Targeted Oxygen Therapy in Adult Intensive Care Unit Patients

Lass Klitgaard, Thomas

DOI (link to publication from Publisher):
10.54337/aau473635979

Publication date:
2022

Document Version
Publisher's PDF, also known as Version of record

TARGETED OXYGEN THERAPY IN ADULT INTENSIVE CARE UNIT PATIENTS

BY
THOMAS LASS KLITGAARD

DISSERTATION SUBMITTED 2022

AALBORG UNIVERSITY
DENMARK

TARGETED OXYGEN THERAPY IN ADULT INTENSIVE CARE UNIT PATIENTS

by

Thomas Lass Klitgaard

AALBORG UNIVERSITY
DENMARK

Dissertation submitted 2022

Dissertation submitted: January 2022

PhD supervisor: Professor Bodil Steen Rasmussen
Department of Anaesthesia and Intensive Care,
Aalborg University
Hospital & Department of Clinical Medicine,
Aalborg University, Aalborg Denmark

Assistant PhD supervisor: Associate Professor Olav Lilleholt Schjørring
Department of Anaesthesia and Intensive Care,
Aalborg University Hospital & Department of Clinical
Medicine, Aalborg University, Aalborg, Denmark

PhD committee: Clinical Associate Professor Helle Damgaard Zacho
Aalborg University, Denmark (chair)

Professor Manu Shankar-Hari
University of Edinburghm, United Kingdom

Professor Jacob Steinmetz
Rigshospitalet and Aarhus University, Denmark

PhD Series: Faculty of Medicine, Aalborg University

Department: Department of Clinical Medicine

ISSN (online): 2246-1302
ISBN (online): 978-87-7573-948-6

Published by:
Aalborg University Press
Kroghstræde 3
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2022

PREFACE

This PhD thesis is focused on oxygen supplementation for critically ill patients admitted to the intensive care unit (ICU) under normobaric conditions (pressure equivalent to that at sea level). It is based on three scientific papers that I completed during my PhD fellowship at the Department of Anaesthesia and Intensive Care at Aalborg University Hospital from 2019-2022. First, the thesis describes the conduct and short-term results of a multicentre, randomised, clinical trial, the 'Handling Oxygenation Targets in the Intensive Care Unit' (HOT-ICU) trial, comparing a lower and a higher oxygenation target in adults acutely admitted to the intensive care unit with acute hypoxaemic respiratory failure. Second, the results of a pre-planned, secondary, Bayesian analysis of the HOT-ICU trial, investigating both the probabilities of different effect sizes, but also the presence of heterogeneous treatment effects are presented. Finally, the results of an updated Cochrane review on higher versus lower oxygenation strategies in the ICU are presented.

Many people have been involved in the work presented in this thesis, and they all deserve my utmost respect and appreciation. Most of all this applies to all the patients and their relatives that were the foundation for the HOT-ICU trial. I would also like to thank all involved staff, be that research or regular ICU nurses, site investigators, or clinicians who have partaken in recruitment of patients and collection of data for the HOT-ICU trial.

I would also like to thank my primary supervisor Professor Bodil Steen Rasmussen, for allowing me to become a part of the department's research unit. The past three years have been some of the most developing, and joyous in my professional career. However, they have also, at times, been some of the most stressful. Thank you for giving me this unique opportunity to experience and learn first-hand about the conduct of state-of-the-art clinical research, and supporting me in this work – especially during the tougher times. The skills I have learned will, I am certain, have a positive impact on my future work as a doctor. In addition, Bodil deserves special acknowledgment for creating a research environment in which one always feels comfortable, competent, and safe.

The first part if this thesis is concerned with the short-term results from the HOT-ICU trial. The trial was designed, and recruitment of patients well underway when I started my PhD-fellowship in May 2019. With little-to-no experience I took over the role as co-ordinating investigator from my assistant supervisor Olav Lilleholt Schjørring, whose PhD thesis most brilliantly describes the initial work and eventual design of the HOT-ICU trial. Olav enthusiastically, and with great care, supported me in this job, and he has truly introduced me to the pleasures of precision and accuracy. Both he and Bodil deserve much credit for inviting me to partake in this scientific

endeavour, and for allowing me to claim my part of the project. The second part of the thesis is concerned with the application of the 'old', but still cutting-edge, statistical computations of Bayesian statistics. For development of the statistical protocol and the subsequent conduct of the secondary analyses of the HOT-ICU trial within this framework, I am grateful to Anders Granholm. He has tirelessly supported me and in a most capable manner answered my many questions concerning the codework and the deeper understandings of the underlaying mathematical principles. Without his support this project would not have succeeded. The last part of this thesis is the based on an updated Cochrane review, for which I have received great guidance and sparring from especially Marija Barbateskovic, the main author of the original review. Marija was always available and quick to answer my many questions in a most competent manner.

Many thanks to all staff involved in the *Collaboration for Research in Intensive Care* (CRIC) and at the *Copenhagen Trial Unit* (CTU), without whom the HOT-ICU trial could not have been completed. A special thanks to our data-manager and computer scientist Janus Engstrøm for his enormous support and monumental effort with the setup and conduct of the HOT-ICU trial, and for always being ready to answer my questions and responding to my many requests.

Thanks to the remaining members of the HOT-ICU trial management committee Anders Perner, Theis Lange, and Jørn Wetterslev for their constant great support and highly qualified feed-back on my scientific work.

I am also grateful to the fantastic staff at the Department of Anaesthesia and Intensive Care's research unit in Aalborg, without whom none of the research we conduct is possible: Stine Rom Vestergaard, Anne Marie Gellert Bunzel, Anne Sofie Broberg Eriksen, Hanne Aaris Mouritsen, Rine Moulvad Siegumfeldt, and Tina Jørgensen. Also, a heartfelt thanks to all the remaining staff in the research unit of Anaesthesia and Intensive Care, and also at the Department of Pulmonary Medicine, for being the foundation of a fun, creative, inspiring, and very enjoyable working environment.

Finally, I wish to express my deepest appreciation and love of my fantastic wife, Tine, for supporting me through the past many years, and throughout my PhD fellowship. Though it at times has been tough, you have always been my solid point of reference and provided me and our children with a safe and steady base. For this I am eternally grateful.

Thomas Lass Klitgaard
Aalborg, January 2022

PAPERS:

I. Lower or Higher Oxygenation Targets for Acute Hypoxemic Respiratory Failure

Schjørring OL*, **Klitgaard TL***, Perner A, Wetterslev J, Lange T, Siegemund M, Bäcklund M, Keus F, Laake JH, Morgan M, Thormar KM, Rosborg SA, Bisgaard J, Erntgaard AES, Lynnerup ASH, Pedersen RL, Crescioli E, Gielstrup TC, Behzadi MT, Poulsen LM, Estrup S, Laigaard JP, Andersen C, Mortensen CB, Brand BA, White J, Jarnvig IL, Møller MH, Quist L, Bestle MH, Schønemann-Lund M, Kamper MK, Hindborg M, Hollinger A, Gebhard CE, Zellweger N, Meyhoff CS, Hjort M, Bech LK, Grøfte T, Bundgaard H, Østergaard LHM, Thyø MA, Hildebrandt T, Uslu B, Sølling CG, Møller-Nielsen N, Brøchner AC, Borup M, Okkonen M, Dieperink W, Pedersen UG, Andreasen AS, Buus L, Aslam TN, Winding RR, Schefold JG, Thorup SB, Iversen SA, Engstrøm J, Kjær MBN, Rasmussen BS for the HOT-ICU Investigators.

New England Journal of Medicine. April 8, 2021;384(14):1301-11.
(Online ahead of print January 20, 2021)
*Shared first authorship
doi: 10.1056/NEJMoa2032510

II. Lower versus higher oxygenation targets in ICU patients with severe hypoxaemia: secondary Bayesian analyses of mortality and heterogeneous treatment effects in the HOT-ICU trial

Klitgaard TL, Schjørring OL, Lange T, Møller MH, Perner A, Rasmussen BS, Granholm A.

British Journal of Anaesthesia. 2021; 128(1):55-64.
(Online ahead of print October 19, 2021)
doi: 10.1016/j.bja.2021.09.010

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

Klitgaard TL, Schjørring OL, Nielsen FM, Meyhoff CS, Perner A, Wetterslev J, Rasmussen BS, Barbateskovic M.

Submitted to the ***Cochrane Database of Systematic Reviews*** on January 12, 2022. Art. No.: CD012631.
doi: 10.1002/14651858.CD012631.pub3

ENGLISH SUMMARY

Oxygen is one of the most prescribed medical drugs world-wide and essential for the proper functioning of the human body's cells and organs. It has been supplied liberally to patients to prevent hypoxaemia and ultimately death. However, oxygen has several known harmful effects and is toxic in high concentrations. An association between hyperoxaemia and increased mortality in acutely ill patients has been found. Yet only a few randomised clinical trials (RCT) in the intensive care unit (ICU) have investigated the issue; some suggesting a benefit of a lower oxygenation strategy, and others suggesting benefit from a higher oxygenation strategy. No RCT yet has been properly designed to evaluate the mortality effects of different oxygenation strategies.

The purpose of this PhD-thesis was to: 1) assess the benefits and harms of two different oxygenation targets in patients with acute hypoxaemic respiratory failure acutely admitted to the ICU; 2) investigate the correlation between organ failure at ICU admission and mortality dependent on oxygenation strategy; and 3) update a systematic review assessing the overall effects regarding targeted oxygenation therapy in ICU patients. The first study was an RCT randomising patients with acute hypoxaemic respiratory failure acutely admitted to the ICU, to either a lower or a higher oxygenation target: the 'Handling Oxygenation Targets in the Intensive Care Unit' (HOT-ICU) trial. It was conducted in 35 ICUs in 7 countries from June 2017 to August 2020, including a total of 2,928 patients. No significant differences in terms of mortality, need for life-support, stay in hospital, or occurrence of serious adverse events (SAE) were found. The second study probabilistically evaluated the mortality effect in the HOT-ICU trial using Bayesian statistical methods, and indicated that large mortality effects were unlikely. However, we saw a potential benefit of the higher oxygenation target in patients with increasing levels of circulatory failure at ICU-admission, resulting in a lower mortality risk in this subset of patients. In the systematic review we identified a total of 16 RCTs on higher versus lower oxygenation strategies, with 6,486 randomised patients. Trials were highly diverse in terms of included patients, duration of intervention, and definitions of such. No differences were found for mortality, number of SAEs or lung injuries, or in self-reported quality of life. The certainty of evidence was low or very low, making firm conclusions difficult.

The findings presented in this thesis have contributed considerably to the evidence concerning targeted oxygen therapy for the adult ICU patient, and demonstrates that if ICU patients' oxygenation generally is targeted within the relative normoxic range, oxygen therapy does not contribute to increased risk of death or SEAs. However, the exact effect is still uncertain and may differ within subgroups, thus additional data is warranted to explore this matter further and provide better treatment.

DANSK RESUMÉ

Ilt er et af de mest anvendte lægemidler på verdensplan og er nødvendig for at vores celler og organer kan fungere optimalt, og har, for at forhindre lave niveauer i blodet, generelt været givet til patienter uden særlige begrænsninger. Som alle andre lægemidler har ilt dog også skadelige bivirkninger og er giftigt i høje koncentrationer. En mulig sammenhæng mellem høje iltniveauer i blodet og overdødelighed er påpeget for akut syge patienter. Blandt patienter på intensiv afdeling er denne sammenhæng dog kun undersøgt i få kliniske studier hvoraf nogle peger på fordel af lavere niveauer af ilttilskud, mens andre peger på fordel af højere.

Formålet med denne ph.d.-afhandling var at: 1) undersøge fordele og ulemper ved to forskellige niveauer af iltning hos voksne patienter med akut hypoksisk lungesvigt (dårlig iltning af blodet grundet lungeskade) akut indlagt på ITA; 2) undersøge sammenhængen mellem graden af organsvigt på indlæggelsestidspunktet på ITA og dødelig afhængigt af iltningsstrategi; og 3) opdatere en systematisk litteratur gennemgang på dette område. Første studie var et lodtrækningsforsøg hvor patienter med akut hypoksisk lungesvigt akut indlagt på ITA blev fordelt til enten et højere eller lavere iltningsmål under intensiv-indlæggelse (HOT-ICU-studiet). Det blev gennemført på 35 intensivafdelinger i 7 lande fra juni 2017 til august 2020, og involverede i alt 2.928 patienter. Der var ingen forskel på dødelighed, behovet for livsunderstøttende behandling, tid på hospital eller forekomsten af alvorlige bivirkninger mellem de to grupper. Andet studie var en Bayesiansk analyse af HOT-ICU-studiet. Resultaterne viste, at store effekter på dødelighed ikke var sandsynlige, men for patienter med kredsløbssvigt på indlæggelsestidspunktet var der muligvis var en fordel ved at stile mod et højere iltniveau. I den systematiske litteraturgennemgang identificerede vi 16 studier vedrørende højere versus lavere iltningsstrategier på ITA med i alt 6.486 patienter. Studierne var meget forskellige både i forhold til hvilke patienter der deltog, hvor lang tid forsøgene løb over og hvordan behandlingerne i studierne var defineret. Vi fandt ingen forskelle i risiko for død, alvorlige bivirkninger, lungeskade eller i livskvalitet. Kvaliteten af evidensen var dog lav og det er derfor vanskeligt at konkludere noget endeligt.

Resultaterne, der præsenteres i denne afhandling, har bidraget væsentligt til den samlede viden om målrettet iltbehandling af voksne intensivpatienter, og viser at såfremt disse patienters iltniveauer holdes i hvad der betragtes som normalområdet, bidrager iltbehandling ikke til øget risiko for død eller alvorlige bivirkninger. Dog er den overordnede effekt stadig usikker og kan variere i udvalgte patientgrupper, hvorfor det fortsat er vigtigt at undersøge dette emne for at kunne levere den bedst mulige patientbehandling.

ABBREVIATIONS

α-level	Risk of type I error (false positive)
ABG	Arterial blood gas
ARDS	Acute respiratory distress syndrome
β-level	Risk of type II error (false negative)
BF	Bayes factor
CI	Confidence interval
CrI	Credibility interval
CPAP	Continuous positive airway pressure
COPD	Chronic obstructive pulmonary disease
DMSC	Data management and safety committee
eCRF	Electronic case report form
ECMO	Extra-corporal membrane oxygenation
EQ-5D-5L	EuroQol 5 dimensions 5 levels questionnaire
EQ-VAS	EuroQol visual analogue scale
FiO_2	Fraction of inspired oxygen
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation
HOT-ICU	Handling Oxygenation Targets in the Intensive Care Unit
HTE	Heterogeneous treatment effects
ICU	Intensive care unit
IMV	Invasive mechanical ventilation

IQR	Inter-quartile range
kPa	Kilo-Pascal
mmHg	Millimetres of mercury
MH fixed	Mantel-Haenszel fixed effect model
MH random	Mantel-Haenszel random effects model
MPV	Major protocol violation
NIV	Non-invasive mechanical ventilation
OR	Odds ratio
PaO_2	Partial pressure of arterial oxygen
PEEP	Positive end-expiratory pressure
RCT	Randomised clinical trial
RD	Risk difference
ROS	Reactive oxygen species
RR	Relative risk
SAE	Serious adverse event
SaO_2	Arterial oxygen saturation
SOFA	Sequential Organ Failure Assessment
SpO_2	Peripheral oxygen saturation
TSA	Trial Sequential Analysis

TABLE OF CONTENTS

'To O'

Capacity of one
Without life be gone
Admire the power
Of what may devour
Double-edged sword
Striking life's chord
Taking and giving
For most living
Too much – too little
Existence hence brittle

TLK, 2022

1. BACKGROUND

1.1. Introduction

Over time, most living beings on earth, humans included, have evolved to become dependent on oxygen for their survival.[1] The discovery of oxygen as a specific compound is generally accredited to the British chemist Joseph Priestley (1733-1804) as he produced oxygen by heating red mercuric oxide on August 1, 1774, and in 1775 was the first to publish an account of the gas.[2] However, it was in fact the Swedish-German chemist Carl Wilhelm Scheele (1742-1786), working independently of Priestley, who, in 1770-1771, was the first to isolate oxygen. However, his results were not published until 1777.[3] After meeting with Priestley in October 1774, the French chemist Antoine-Laurent de Lavoisier (1743-1794) repeated Priestley's experiments, and went on to explain the nature of the gas, naming it 'oxygen'.[4,5] Of note, Scheele had written to Lavoisier on September 30, 1774 with his experiences on the matter, but neither Scheele nor Priestley were accredited by Lavoisier in his publications. Though the isolation of oxygen and its role in respiration was clarified in the late 18th century, understandings of respiratory physiology dates back to the 13th century where already in 1250, the Syrian physician Ibn al-Nafis described the pulmonary circulation. He thus predated William Harvey's much-famed publication on this subject from 1628, the 'Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus', in which Harvey explains the principles of circulation of blood in the body.[6] Even though the potential for oxygen's use for medicinal purposes was quickly recognised, among others by Priestley,[2] scientists were also aware of the potential toxic side effects of oxygen. Notably, the results from the experiments of French physiologist Paul Bert (1833-1886) on oxygen toxicity were published in 1878:[7] high levels of oxygen (hyperoxia) would lead to convulsions and death, now called the 'Paul Bert effect' when referring to cerebral oxygen toxicity. Later in 1899, the Scottish pathologist James Lorrain Smith (1862-1931) demonstrated the toxic effects upon the lungs of inhaling high fractions of normobaric oxygen over prolonged periods of time;[8] aptly named the 'Lorrain Smith effect'. These findings of oxygen toxicity have stood the test of time, and are still highly relevant to both patients and clinicians in modern health care systems, as further elaborated below. On the other hand, low levels of oxygen in the blood (hypoxaemia) are also problematic as they may result tissue hypoxia leading to cellular dysfunction, organ failure, and ultimately death, thus prompting the question: *what is the optimum level of oxygenation to target*?

In most countries worldwide, oxygen is classified as a medical drug, and is thus subjected to the same regulations as other medical drugs. Consequently, oxygen must be prescribed to patients with both beneficial and potential harmful effects in mind, with a target of 94-98% arterial saturation being recommended by the latest

guidelines for acutely ill patients in general.[9–11] Arguments for a lower oxygenation target range (oxygen saturation of 92-96%) have been raised, especially emphasising a lower risk for above-normal levels of blood oxygen content (i.e. hyperoxaemia).[12] Interestingly, current guidelines do not pertain to patients admitted to the intensive care unit, due to limited data at the time of their publications. Specific recommendations for ventilatory strategies for patients with acute respiratory distress syndrome (ARDS) admitted to the ICU do exist, but without any detailed recommendations regarding oxygenation.[13]

It must at any rate be plain to any candid mind that oxygen is a real though as yet not very well understood therapeutic power. It is the bitterest sarcasm on our respectable and conventional system of therapeutics that nothing like a concerted *effort has yet been made by competent and credible men in England to settle what the true functions of so powerful a therapeutic weapon may be*

– Francis Edmund Anstie (1871)[14]

Despite oxygen being known to the medical profession for more than 250 years, the optimal oxygenation strategy is still a matter of debate, and is regrettably still not clearly defined.

1.2. Oxygen toxicity

In aerobic (from Ancient Greek: aéros = *air* + bíos = *life*) metabolism, molecular dioxygen (O_2) is reduced in the mitochondrion in order to produce adenosine triphosphate via oxidative phosphorylation. In this process, a number of highly reactive intermediate metabolites are produced: the reactive oxygen species (ROS), being the superoxide anion ($^{\bullet}O_2^-$), hydroxyl radical ($^{\bullet}OH$), and hydrogen peroxide (H_2O_2).[15] These biproducts of mitochondrial energy production may react with lipids, proteins, and DNA/RNA, leading to cellular damage, and eventually destruction.[16] Under normal conditions this is counterbalanced by the body's endogenous anti-oxidant system, but may be overwhelmed in the case of hyperoxia, and especially

within the lungs as this organ is subjected to the highest concentration of oxygen.[17,18] The excessive ROS production is thus believed to be the main mediator of pulmonary toxicity, and cellular dysfunction generally.[16,19,20] Since the early realisation that oxygen in higher concentrations is toxic, a range of harmful side-effects has been described.[21] In the following a short overview of the most important will be presented.

Reports in the first half of the 20th century on oxygen tolerability in healthy subjects demonstrated that relatively short-term exposures (up to a few hours) to high oxygen concentration, at above-normal pressure, would result in neurological symptoms, including seizures similar to the findings of Paul Bert, with clearly decreased tolerance with increasing pressures of the inhaled oxygen.[22] In 1970, Barber et al. demonstrated the toxicity of prolonged inhalation of pure oxygen on the lungs with impaired gas exchange, increased intra-pulmonary shunt, decreased pulmonary compliance, and increased lung weight.[23] These findings were coalesced most brilliantly in the review of oxygen toxicity by Clark and Lambertsen in 1971.[17] Additional pulmonary toxic effects include absorption atelectases (leading to right-to-left shunting), consolidation, congestion, inflammation, and fibrin formation, among others.[17,24] Hyperoxaemia have additional adverse effects: vasoconstriction resulting in paradoxical tissue hypo-perfusion and reduced and inadequate oxygen levels (hypoxia) e.g. in the brain and myocardium; ocular and retinal damage, including blindness – especially in the new-born; testicular damage; and erythrocyte haemolysis.[17,25–27] At fractions of inspired oxygen (FiO_2) approaching 1.00, survival times of a range of animals were summarised in the review by Clark and Lambertsen, with approximately 39 hours in a dog to more than 1600 hours in a frog, and with decreasing tolerability with increasing age.[17] However, most information at that time was from either animal studies or studies on healthy volunteers and not in the critically ill (i.e. admitted to the ICU). More recently, the level of toxicity in humans is suggested to be dependent on both the level of oxygenation (and partial pressure applied) and the duration of exposure.[28]

What is there that is not poison? All things are poison, and nothing (is) without poison. Solely the dose determines that a thing is not a poison

Phillipus Theophrastus Aureolus Bombastus von Hohenheim ('Paracelsus') (1493/94–1541), from his 'Third defence'[29]

The interest in oxygen therapy in critically ill patients (i.e. those admitted to the ICU) has increased in recent years, with a number of systematic reviews and meta-analyses on this matter being published.[30–33] However, the reviews report conflicting results, as some find that supplemental oxygen potentially is harmful or at least not beneficial,[30,31] whilst others found the evidence insufficient to support either beneficial or harmful effects.[32,33] In addition, discrepancies exists when comparing actual measured oxygenation levels with doctors' preferred oxygenation targets.[34] A more conservative oxygenation strategy, targeting partial pressure of arterial oxygen (PaO_2) levels in the 'low-normal' to 'subnormal' range (7.3-10.7 kPa), has been proposed in an attempt to mitigate hyperoxia-induced lung injury; dubbed 'permissive hypoxaemia'.[35–37] The superiority of this concept as compared to normal blood oxygen levels (normoxia) remains yet unproven.[38]

1.3. Oxygen use in the intensive care unit

Normal pulmonary gas exchange can be impaired by many different factors, both pulmonary (e.g. pneumonia, lung contusions, lung cancer, or inhalation of noxious gasses) and extra-pulmonary (e.g. sepsis, and multiple trauma).[39,40] This impairment, if severe enough, will lead to hypoxaemia (i.e. hypoxaemic respiratory failure) and may lead to admission to an ICU. To prevent or treat hypoxaemia and subsequent tissue hypoxia, patients admitted to the ICU are prescribed supplemental oxygen. The level of hypoxaemic respiratory failure is generally evaluated by means of the ratio between arterial partial pressure of oxygen (PaO_2) and the fraction of inspired oxygen (FiO_2), the 'PaO_2/FiO_2 ratio', with lower ratios denoting increasing severity of lung failure. It also plays a central role in the definition of acute respiratory distress syndrome (ARDS), as it dictates the perceived level of severity.[41] Though most probably a continuum of respiratory failure, the level of severity, has for the sake of simplicity, been categorised as being either 'mild' (26.7 kPa < PaO_2/FiO_2 ≤ 40 kPa), 'moderate' (13.3 kPa < PaO_2/FiO_2 ≤ 26.7 kPa), or 'severe' (PaO_2/FiO_2 ≤ 13.3 kPa). However, the relationship between FiO_2 and the PaO_2/FiO_2 ratio is not universally linear, plus the ratio may be also affected by intrapulmonary shunting, haemoglobin concentration, and SaO_2, thus complicating the interpretation of the ratio as a direct measure of pulmonary failure.[42–45]

Oxygen is most commonly supplied by means of inhalation, either by open oxygen supplementation systems (e.g. nasal catheters, face-mask, high-flow systems, etc.) or closed oxygenation systems (e.g. mask/helmet continuous positive airway pressure [CPAP], non-invasive mechanical ventilation [NIV], or invasive mechanical ventilation [IMV]). In certain circumstances, it can also be provided by means of extracorporeal membrane oxygenation (ECMO), e.g. in case of severe respiratory failure or during cardio-pulmonary surgery. Lastly, oxygen may be provided by

inhalation at hyperbaric conditions, where the patient is placed in a special pressure chamber and the pressure is increased above ambient pressure (e.g. 2-3 times normal). As stated earlier, the focus of this thesis will be on the supplementation of normobaric oxygen via inhalation. The 'standard' approach to oxygen supplementation has often resulted in patients being hyperoxaemic,[28,46,55–64,47–54] however, both hypo- and hyperoxaemia have been associated with increased mortality in several observational studies,[46,50,51,55–61] whilst not in others.[48,54,62–64] The design and planning of the HOT-ICU trial began already in 2015, before the publication of any major RCTs on this matter, but at the time of trial initiation, there was increasing attention in the intensive care community to the potential harmful effects of oxygen supplementation, and a general notion that lower levels of oxygen appeared to be the better option.

When I started my PhD fellowship in May 2019 and joined the HOT-ICU team, only two major randomised clinical trials,[65,66] several small-scale pilot trials,[67–71] low-powered randomised trials,[72–74] and few before-and-after-trials[75–78] on oxygenation strategies in adult ICU patients had been published. I will briefly introduce the primary RCTs below.

Girardis et al. reported in October 2016 the results from their 'Normal Oxygenation Versus Hyperoxia in the Intensive Care Unit' (OXYGEN-ICU) trial.[65] This was an Italian, single-centre trial, conducted in Modena from March 2010 to October 2012, involving 480 patients admitted to the medical ICU with an expected ICU stay of at least 72 hours. Patients were randomised to either a 'conservative' oxygenation strategy (PaO_2 9.3-13.3 kPa or SpO_2 94-98%) or a 'conventional' oxygenation strategy (FiO_2 ≥0.40, PaO_2 ≤20 kPa, or SpO_2 97-100%) for the duration if ICU admission. Due to a massive earthquake destroying the hospital in May 2012, the trial was prematurely stopped before reaching the planned 660 patients. At baseline, roughly two thirds of patients in this trial were mechanically ventilated, but only a little more than half of the included patients had respiratory failure according to the investigators. Unfortunately, no baseline measures of oxygenation were reported, therefore comparison of severity respiratory failure to other trials is not possible. The investigators achieved only a minimal, albeit statistically significant, separation in oxygenation between the two groups: median FiO_2 0.36 versus 0.39, and median PaO_2 11.6 kPa versus 13.6 kPa in the conservative and conventional group, respectively. By the end of the trial 11.6% percent in the conservative group and 20.2% in the conventional group had died in the ICU; RR 0.57 (95% CI 0.37-0.90, p = 0.01). However, the final analysis was conducted in a modified (unjustified and not pre-specified) intention-to-treat cohort.

Asfar et al. reported in February 2017 the results from their 'Hyperoxia and Hypertonic Saline in Patients with Septic Shock (HYPERS2S)' trial.[66] The trial was conducted from November 2012 and June 2014, and was a French, multicentre, 2x2 factorial trial conducted in adult patients with septic shock with a PaO_2/FiO_2 ratio

≥13.3 kPa. Patients were randomised to either 'hyperoxia' (FiO_2 1.0) or 'normoxia' (SaO_2 88-95%) for the first 24 hours of ICU admission after randomisation. Patients were also randomised to intravenous infusions with either isotonic or hypertonic saline. After the second interim analysis, the trial was prematurely halted due to safety reasons and lack of benefit (in both intervention domains). A total of 442 of the planned 800 patients were randomised. Patients had a median PaO_2/FiO_2 ratio of 26.5 kPa at baseline, equivalent to mild-moderate degrees of ARDS. The investigators did not find any significant difference in all-cause mortality, but did find an increase in the proportion of patients with one or more serious adverse event (SAE) in the hyperoxia-group.

In 2018, Chu et al. published the 'IOTA study', a systematic review and meta-analysis, investigating the mortality effect of higher versus lower oxygenation strategies in a mixed cohort of acutely ill patients.[79] The authors concluded that a higher oxygenation strategy was potentially harmful to patients, and the study resulted in a rapid recommendation on oxygen supplementation for the acutely ill patient.[80] These findings were corroborated by the, at that time, most recent Cochrane review on oxygenation strategies in the ICU, conducted during the inclusion period of the HOT-ICU trial.[81] However, the certainty of evidence in the Cochrane review, in opposition to what was concluded IOTA study, was deemed very low due to both risks of bias, large differences in the applied interventions when comparing the included trials, and an overall small sample size.

Since the start of the HOT-ICU trial, the results from three additional large-scale RCTs have been published.[82–84] The findings of these RCTs, and the previously mentioned, in relation to the findings of the HOT-ICU trial, will be discussed later in this thesis (see section **Fejl! Henvisningskilde ikke fundet.**).

1.4. A brief introduction to Bayesian statistics

As a substantial part of this thesis pertains to the conduct of Bayesian analyses, I will in the following section present a brief introduction to the principles of Bayesians statistics. The intention is not to provide an in-depth understanding of the complex mathematics behind, but instead to offer an overview of the general principles, such that the findings in Paper II may be more easily understood and appreciated.

1.4.1. 'Classical' and 'Bayesian' statistics

The probability of an event, or outcome, is described as a proportion between 0 and 1, equivalent to either never occurring (i.e. 0) or always (i.e. 1), or somewhere between. This probability is often expressed in percentages (i.e. 0 to 100%). There are several statistical frameworks in which this matter may be approached: *frequentist* or *Bayesian* being two of such (*law of likelihood* being a third, which in turn is utilised within the Bayesian framework). In classical, frequentist statistics parameters (or variables of interest) are treated as unknown but fixed quantities. They are estimated by sampling from various sample distributions, and the estimate is described (e.g. by the median or mean value) together with the uncertainty of this estimate (e.g. interquartile range [IQR] or standard deviation [SD]). However, the parameter of interest is still considered fixed. Frequentist analysis will ultimately end up testing the 'null-hypothesis'; often formulated as there being no difference in outcome between two interventions or groups of participants. The analysis yields a *p-value* (the probability to obtain a test statistic as extreme or more extreme, as represented by one's sample, given that the null-hypothesis is true) and a *confidence interval* (CI). The CI is the range of values likely to contain the unknown parameter's *true* value, and if sampling were to be repeated a proportion of the confidence intervals produced would contain the true parameter value. This proportion is equivalent to the *confidence level*. However, this cannot be inversely interpreted as the true value will be contained, with 95% probability, within the 95% CI calculated based on one's specific sample. A trial's probability to incorrectly reject the null-hypothesis (a type 1 error = false positive), and thus finding a significant difference when non exists is denoted by the *significance level* or α-level. Also, the more inferences are made the greater the risk of spurious findings, if not properly accounted for.[85] Both the significance level of the p-value, often set at 5%, and the corresponding 95% CI are arbitrarily chosen. On a side-note, it was Ronald Aylmer Fisher (1890-1962), by many considered the father of modern frequentist statistics, who developed the statistics of the p-value and suggested the 5%-level for statistical significance most use today.[86] The use (and misuse) and proper interpretation of the p-value has been heavily debated in recent years, with some arguing for the total abandonment of its use, whilst others argue for a more nuanced and cautious interpretation (and avoidance of dichotomisation) of the p-value when informing clinical practice.[87] Also, complete abandonment of the terms "statistically significant" has been proposed by the American Statistical Association, and they instead emphasise that one must be modest, sceptical, and open when interpretating of one's findings.[88]

In the Bayesian approach, the parameter of interest is considered uncertain and described with a probability distribution. Bayesian statistics and Bayes' theorem (or rule) is named after the English statistician Thomas Bayes (1701-1761), although his formulations were first published posthumously in 1763 by Richard Price (after having edited the thesis himself).[89] However, the broader formulations and

expansions of these principles are mainly accredited to the French mathematician Pierre-Simon Laplace (1749-1827) and his 'Essai philosophique sur les probabilités' (1814).[90] Laplace never used the phrase 'Bayesian statistics', but instead coined 'inverse probability' (which is now considered obsolete). By using the foundation laid out by Bayes, Laplace would formulate the general principles of (what is now called) Bayesian statistics, and use known quantities to estimate probabilities for unknown parameters. It was in fact Fisher who first to have used the term 'Bayesian', but as Fisher was never a fan of the concept he had used it in a condescending manner.[91]

Thomas Bayes predominantly worked on 'conditional probability' and formulated the mathematical rule to calculate the probability of an event (A) *given the occurrence of another* (B), written as P(A|B). This can be derived as follows: the joint probability of the events A and B (P(A ∩ B)) occurring together is equal to the probability of event A occurring multiplied with the probability that event B occurs *given* the occurrence of event A:

$$P(A \cap B) = P(A)P(B|A)$$

The joint probability of events A and B is also equal to the probability of event B occurring times the probability of event A occurring *given* the occurrence of event B:

$$P(A \cap B) = P(B)P(A|B)$$

These two can then be equated:

$$P(B)P(A|B) = P(A)P(B|A)$$

and may be rearranged to produce Bayes' rule:

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

which is often rewritten:

$$P(H|E) = \frac{P(H)P(E|H)}{P(E)}$$

Where P(H|E) denotes the probability of the outcome *H* (or hypothesis) given the data *E* (or evidence), and P(H) and P(E) are *prior* probabilities of observing the hypothesis or data, respectively. P(E|H) is the conditional probability of observing the data given the hypothesis, often named the *likelihood function*. To ease interpretation, this may be further simplified to:

$$Posterior \propto Likelihood\ x\ Prior$$

This means that the *posterior probability distribution* (or 'posterior') of the observed event is proportional ($\propto$) to the *likelihood function* (defined by the data) times the *prior probability distribution* (or 'prior'). Typically, the posterior is described with a median or mean value and a 95% *credibility interval* (CrI). The CrI is simply an integral of the posterior probability distribution and may be interpreted as the range wherein the true value with 95% probability will be, given our data, the prior, and the statistical model used. In order to conduct a Bayesian analysis, one therefore needs to define the prior(s), and this is optimally done *before* the results of a trial are investigated (*a priori*), but may be formulated post-hoc.[92,93] There is a vast variety of possibilities for defining a prior, and the choice may depend on one's previous knowledge of the question at hand, or may simply contain a range of probabilities that one may assume plausible, e.g. that the distribution belongs to a normal distribution with a certain mean and standard deviation. Even if no information of the studied matter exists, one may be better off by instituting an informed 'guess' (e.g. based on clinical experience, expert opinion, or information from comparable fields of research) than using a *flat prior* (equal to a uniform distribution), being one that assigns equal probability to all outcomes of the parameter of interest.[94,95] Priors may, to a lesser or a greater extent, favour the occurrence of the studied event (e.g. be pessimistic or optimistic), be sceptical of large event variations, etc. If one implements minimally informative priors (only containing minimal information regarding assumptions of the variable in question) the results of the Bayesian analysis will very closely resemble that of the classical frequentist analysis, but without the constraints of the p-value and corresponding confidence interval.[93] The more information contained in the prior, the more influence it will have on the posterior. Likewise, increasing the size of the data provided will also have increasing effect on the posterior, e.g. when combining a weakly informative prior with data from a large trial, the data from the trial will generally overwhelm the prior and dominate the posterior. Thus, precision of the posterior is jointly dependent on the information in both the prior and data provided – and naturally the specified statistical model (but this is no different to the frequentist approach). The posterior distribution of one trial (or several) may also be implemented as the prior in another trial, thus updating the joint information on the subject. One example of this could be implementing information from a meta-analysis of previous trials and using this information as the prior when estimating the posterior based on results from a new trial (typically as a sensitivity analysis).[96] As the posterior is the final result from the Bayesian analysis, it is possible to ascertain a multitude of different integrals representing probabilities of various outcomes. For example, it is possible to calculate the probability that the parameter of interest is above (or below) a certain threshold, or within a specified range. Such integrations do not function as new tests and are therefore not beset with the issues of multiple testing that plague frequentist statistics.

Bayesian statistics is sometimes accused of being subjective due to the explicit definitions of the incorporated prior(s). However, the frequentist and Bayesian schools of thought are both based on several assumptions and subjective judgements. For instance, in frequentist statistics, assumptions of the distribution and properties of the data (e.g. linear correlation of parameters) have to be made, and the interpretation of the results from any analysis must be interpreted with these assumptions in mind. Often these assumptions are not explicitly reported when conducting frequentist analyses, but may only be vaguely described in the methods sections.[97] In Bayesian analysis, all assumptions regarding model parameters and priors must be presented,[92,98] thus making this approach transparent to the reader. Though no official reporting guideline exists for Bayesian analyses, Zampieri et al. have recently proposed such a guide for the conduct and reporting of Bayesian analyses.[92] Great care must be employed when designing priors, and it is advisable to implement a range of such, with proper justification, to assess the impact of prior-specifications on the final results (i.e. sensitivity analyses).[93]

Even though the basis for Bayesian statistics is by no means new, conduct of Bayesian analyses requires great computational abilities (and substantial computer power), and entails sophisticated statistical analyses. These aspects may limit its use. Also, statistics can sometimes be regarded as a world of *either/or*: either one is a classical frequentist statistician or a Bayesian. However, the two statistical approaches may complement one another as has been the case in several large scale ICU trials.[99–108] Recently, Bayesian statistics have also been used as the primary statistical framework in a number of publications from the REMAP-CAP trial setup (Randomised, Embedded, Multi-factorial, Adaptive Platform Trial for Community-Acquired Pneumonia).[109–111] The implementation of the Bayesian methods into clinical research practice has only been done within the last few decades, but with the advent of more powerful computers, it is reasonable to expect that this approach to data will become more common-place in the future.

1.5. Appraising the effect of an intervention

In modern medical research, the randomised clinical trial (RCT) is viewed as the gold standard of testing the causal effect of interventions.[112] By randomising patients to one of two (or more interventions) any imbalances of baseline characteristics, potentially leading to bias (selection bias), will be minimised if done properly – especially if stratification for important factors is implemented.[113] As per the Consolidated Standards of Reporting Trials (CONSORT) statement first published in 1996, and updated latest in 2010, reporting of baseline demographics is customary, to demonstrate the effectiveness of the randomisation process.[114,115] However,

testing of baseline differences is not recommended, as any differences are a product of chance, and testing can be misleading.[116]

RCTs are often divided into two main categories: *explanatory* and *pragmatic* trials, as coined by Schwartz and Lellouch in 1967.[117] The former testing the effect of an intervention under optimal conditions, whilst the latter does so under routine-practice conditions, potentially leading to a greater level of external validity. The results of a pragmatic trial can thus be more easily generalised to 'real-world conditions' (depending on the definition of such), but due to higher degrees of heterogeneity among studied subjects larger sample sizes are needed.[118] Ultimately, the pragmatic RCT is '*designed for the primary purpose of informing decision-makers regarding the comparative balance of benefits, burdens and risks of a biomedical or behavioural health intervention at the individual or population level*'.[119] The goal of the RCT is often to demonstrate the superiority of an intervention in comparison to a control condition, being e.g. either an active treatment or placebo.[120] This concept is termed a *superiority* trial. If the premise is to demonstrate equipoise of the two approaches the trial is an *equivalence* trial. Lastly, if the purpose is to demonstrate that one intervention is no worse (importantly at a pre-defined level) than another, it is a *non-inferiority* trial. Choice of trial design depends on the research question (hypothesis) at hand, and this choice has consequences for calculations of the required sample size, analyses of data, and ultimately interpretation and reporting of the trial's results.[115,121] However, the 'classical' RCT has several other limitations to consider: oversimplification of intervention design; exaggerated intervention effects to inform sample size calculations, consequently with an embedded requirement for large effects sizes to appraise benefits or harms; and lack of flexibility, to mention some of the most important.[122]

Typically, a hypothesis is stated as a difference in outcome(s) given two treatments. In the case of the HOT-ICU trial, that '*targeting a lower oxygenation target as compared to a higher target, in acutely ill patients in the ICU with hypoxaemic respiratory failure, would result in an absolute reduction in 90-day all-cause mortality of 5 percentage points (i.e. from 25% to 20%), equivalent to a relative risk of 0.80*'.[123,124] The HOT-ICU trial is thus by definition a superiority trial, as the intervention (the lower oxygenation target) is assumed to be superior to the control condition (the higher oxygenation target). In this specific case the null-hypothesis may be formulated as '*the mortality effect of a lower oxygenation target is no different to that of a higher oxygenation target*', and it is this hypothesis which, in turn, is tested statistically once all data have been collected when employing a frequentist approach. The object of the test is to reject the null-hypothesis – with a certain level of confidence, i.e. at the level of the a-priori defined level of statistical significance. Thus, either providing a statistically significant or insignificant result. However, the inability to reject the null-hypothesis, being no significant difference between the two treatments, is not equivalent to proving evidence of identical levels of treatment effects – i.e. '*absence of evidence is not evidence of absence*'.[125] Besides

evaluating statistical significance, one must as a clinician also appraise the level of clinical significance, i.e. the threshold of a meaningful clinical difference of an outcome. A difference may very well be statistically significant but hold no clinical importance. This depends on the investigated outcome. However, in trials based on frequentist inference, detection of smaller (clinically relevant) differences requires larger sample sizes which may cause great difficulties in trial conduct. It is also important to evaluate the point estimate and the limits of the corresponding interval of uncertainty, as effect sizes pointing in the opposite direction than the point estimate may exist and be clinically relevant. However, one must remember that values closer to the point estimate are more compatible with the data (under the assumptions of the trial) than those further away, and this fact should be carefully considered when interpreting one's findings.[87]

The advent of Bayesian statistics has challenged this classical approach to conducting and interpreting clinical trials. In this setting, no sample size is calculated beforehand. Instead, levels of clinical equipoise, superiority, inferiority, or futility are estimated prior to start of the trial. By doing frequent interim analyses, the result of the trial (being the posterior probability distribution for the pre-defined outcome), can be updated correspondingly using the available trial data, and the trial can continue until either of the specified levels of effect has been reached.[110] This may, for example, result in early termination of a trial in the case of clinical superiority of an intervention, as was the case in the REMAP-CAP IL-6 inhibitor trial.[109] However, a more detailed discussion of the aspects of clinical trial conduct based on Bayesian inference models, including the rapidly developing concept of adaptive platform trials, is beyond the scope of this thesis and will not be dealt with further.

Findings of any RCT may deviate from the 'true' result due to a number of reasons. The trial may be *confounded*, that is being affected by exterior factors that are not accounted for in the trial's design, for instance inability to blind an intervention. Chance can also play an important role, especially if the sample size is very small. Irrespective of the statistical framework, the results may also deviate due to systematic errors in the methodology applied – i.e. bias.[126] The extent of bias should be judged systematically, and when considering RCTs the Cochrane Risk of Bias tool proposed by Higgins et al. in 2011, updated in 2019 to version 2 (RoB-2), is the preferred instrument.[127,128] With proper trial design, the aforementioned sources of error can, to a large extent, be reduced: e.g. by increasing the sample size the effect of chance will be mitigated, and sound trial design will minimise bias. Such aspects are essential to appraise when conducting systematic reviewing, as it may influence the level of confidence one may put in the results from a given RCT, as per the current Grading of Recommendations Assessment, Development and Evaluation (GRADE) assessment guidelines, and ultimately the overall findings of a review.[129] Systematic reviews of all existing RCTs examining a certain clinical question, including properly conducted meta-analyses, with critical appraisal of the certainty of evidence (i.e. GRADE assessment), is considered the highest level of evidence.[130] It is of great

importance to gather all available evidence when informing clinical decision making, but as with RCTs, the conduct of systematic reviews and meta-analyses must be performed with much care and rigor, in order not to draw false conclusions regarding effects.[131] This entails pre-planning of proper research question(s), selection criteria for trials, search strategy, outcomes investigated, analyses, etc., for example by publication of a protocol prior to starting the review process.[126]

2. AIMS AND HYPOTHESES

2.1. The Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial (Paper I)

The aim of the first study (Paper I – Appendix A)[132] was to evaluate the benefits and harms of two oxygenation targets: a lower (PaO_2 of 8 kPa) versus a higher target (PaO_2 of 12 kPa), in patients acutely admitted to the ICU with acute hypoxaemic respiratory failure. The hypothesis was that by targeting a lower oxygenation target, as compared to a higher, the 90-day all-cause mortality would be reduced by 5 percentage points, equivalent to a relative risk reduction of 20%.

2.2. Bayesian and heterogeneity of treatment effect analyses of the HOT-ICU trial (Paper II)

The aim of the second study (Paper II – Appendix B)[133] was to apply Bayesian statistical analysis techniques to probabilistically assess the mortality effects of a lower oxygenation target versus a higher on the primary outcome of the HOT-ICU trial. In addition, the study aimed to evaluate the probabilities of a range of effect sizes, and based on pre-specified baseline variables, to explore the presence of heterogeneous treatment effects on all-cause mortality. As in the first study, the hypothesis was that a lower oxygenation target would reduce mortality as compared to a higher oxygenation target, and in addition that patients with increasing degrees of organ dysfunction (measured by a panel of baseline characteristics) would have increasing benefit of a lower oxygenation target as compared to a higher.

2.3. Updated Cochrane review: Higher versus lower fractions of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (Paper III)

The aim of the third study (Paper III – Appendix C)[134] was to systematically review and update the assessment of the benefits and harms of higher versus lower FiO_2 or levels of arterial oxygenation (i.e. oxygenation strategies) in adult ICU patients. For the purpose of this review, no a-priori hypotheses were formulated.

3. METHODS

3.1. The HOT-ICU trial (Paper I)

3.1.1. Trial design

The HOT-ICU trial was a multicentre, pragmatic, investigator-initiated, stratified, parallel-group trial of two oxygenation targets. The trial was prospectively registered (ClinicalTrials.gov identifier: NCT03174002; EudraCT number: 2017-000632-34), and both the trial's protocol and statistical analysis plan were published prior to randomisation of the last patient.[123,124] No noteworthy changes were made to the trial protocol, except a minor change in the statistical reporting of the pre-planned secondary analysis of the primary outcome due to the nature of the obtained data (see section 3.1.8).

3.1.2. Eligibility criteria

Patients fulfilling the following inclusion criteria within 12 hours of ICU admission were screened for inclusion: 18 years or older; acutely admitted to the ICU; receiving at least 10 litres of oxygen in an open oxygen supplementation system or an FiO_2 of at least 0.50 in a closed oxygen supplementation system (IMV, NIV, or CPAP); expected oxygen supplementation in the ICU for at least 24 hours; and a functioning arterial catheter in place for frequent blood sampling.

Patients fulfilling one or more of the following criteria were excluded: cannot be randomised within 12 hours of ICU admission; receiving chronic mechanical ventilation (IMV, continuous NIV, or continuous CPAP) for any reason; use of supplementary home oxygen; previous bleomycin treatment; organ transplant planned or conducted during index admission; withdrawal from active therapy or brain death deemed imminent; pregnancy; poisoning with either carbon monoxide, cyanide, or paraquat; methaemoglobinaemia; sickle cell disease; hyperbaric oxygen treatment; consent not obtainable according to national regulations; or previously randomised into the HOT-ICU trial. Detailed definitions of inclusion and exclusion criteria are available elsewhere.[123,132]

3.1.3. Randomisation

Included patients were randomised 1:1 using centralised randomisation with a computer-generated concealed assignment sequence with permuted blocks of varying sizes. Inclusion was stratified according to including trial site, the presence or absence of chronic obstructive pulmonary disease (COPD), and the presence or absence of active haematologic cancer. Detailed definitions of stratification variables are provided in the supplement to Paper I (Appendix A).[132]

3.1.4. Interventions

Patients were randomised to either a lower oxygenation target (PaO_2 of 8 kPa), or a higher oxygenation target (PaO_2 of 12 kPa), and were to adhere to this oxygenation target for their entire ICU admission, including any ICU re-admissions, for up to 90 days after randomisation. The lower oxygenation group was defined as the intervention group and the higher oxygenation group as the control group. Treating clinicians were to titrate the FiO_2 between 0.21 and 1.00 in order to obtain the allocated oxygenation target. As the trial was pragmatic in its design, all other treatments, including (but not limited to) choice of ventilator strategy, use of proning, use of inhaled vasodilators, and selection of oxygen supplementation device/system were at the discretion of the treating physician. The use of additional oxygen supplementation during e.g. procedures in the ICU, as well as during surgery, transportation (in-hospital or between hospitals), or radiological examinations was also at the discretion of the treating clinician. The use of an FiO_2 of 1.00 prior to or during airway suctioning or intubation, if not necessary to reach the allocated oxygenation target, was discouraged. However, if this was not deemed possible by the treating clinician, pre-oxygenation was advised to be for maximum duration of one minute prior to endotracheal suctioning, and maximum three minutes prior to intubation. Deviations above the allocated oxygenation target if the FiO_2 was 0.21, and deviations below the target if the FiO_2 was 1.00, were allowed. However, the maintenance of the assigned oxygenation target was requested whenever possible.

3.1.5. Trial outcomes

The primary outcome was 90-day all-cause mortality. Secondary outcomes were: percentage of days alive without life-support within 90 days of randomisation (being use of respiratory support, renal replacement therapy, or circulatory support); percentage of days alive out of hospital within 90 days of randomisation; and the number of patients with one or more SAE in the ICU within 90 days of randomisation.

SAEs were defined as new episodes of shock (plasma lactate concentration ≥ 2 mmol/l and continuous infusion of either a vasopressor or an inotrope), intestinal ischaemia, cerebral ischaemia, or cardiac ischaemia. Details on SAE definitions are provided elsewhere.[123,132]

Additional pre-defined, secondary outcomes not reported in Paper I include: one-year all-cause mortality; health related quality of life measured using the EuroQol 5 dimensions 5 levels questionnaire (EQ-5D-5L) and EuroQol visual analogue scale[135] one year after randomisation; cognitive functioning at one year after randomisation using the Repeatable Battery for the Assessment of Neuropsychological Status at selected sites;[136] pulmonary functioning one year after randomisation using whole body plethysmography at selected sites; and a health economic analysis. These will be provided in separate publications.[124,137]

3.1.6. Sub-groups

The primary outcome was considered in the following subgroups based on baseline characteristics: shock at randomisation; receiving IMV; type of ICU admission (medical/elective surgical/acute surgical); known COPD; acute traumatic brain injury; cardiac arrest <24 hours prior to randomisation; active haematological malignancy; ARDS; and oxygen supplementation through a closed system at randomisation according to baseline PaO_2/FiO_2 ratio (<13.3 kPa; ≥13.3 to < 26.7 kPa; ≥26.7 to < 40.0 kPa; and ≥40.0 kPa). The latter three were planned to be published separately, and results are not presented in this thesis.[124]

3.1.7. Oxygenation measures

The highest and lowest PaO_2 in pre-defined 12-hours intervals were registered: from 06:00 to 18:00 and from 18:00 to 06:00, corresponding to the working day in a Danish ICU. The concomitant SaO_2 and FiO_2 in both time intervals were also registered, thus providing up to four daily registrations of oxygenation parameters for each patient during ICU admission.

3.1.8. Statistical analyses

All statistical analyses were performed according to the published statistical analysis plan[124] and conducted in the intention-to-treat cohort, that included all patients

randomised except for whom consent was withdrawn or unobtainable.[138] Statistical assessment of all reported outcomes was blinded to group allocation.

Sample size calculation

The mortality in the control group (higher oxygenation group) was estimated to 25% based on several sources: a multicentre, observational cohort study;[55] previous findings of a two Scandinavian cohort studies on severe sepsis or septic shock;[50] and a cohort study on risk factors for gastrointestinal bleeding.[139] To detect or reject a true relative risk reduction of 20% in 90-day all-cause mortality (equivalent to an absolute risk reduction of 5 percentage points) from the estimated 25% risk in the control group, with a two-sided α-level of 5%, a β-level of 10% (equivalent to a power of 90%), 2,928 patients were required.

Statistical significance

For the primary outcome of 90-day all-cause mortality, a 95% confidence interval (CI) not including 1.00 for the risk ratio (RR) equivalent to a two-sided p-value less than 0.05, was considered statistically significant.

CIs for the secondary outcomes were adjusted based on the five-step procedure as suggested by Jakobsen et al.[140] With a total of seven secondary outcomes, the adjusted p-values would be below 0.0125 to yield statistical significance, corresponding to adjusted CIs of 98.75%, in order to preserve a family wise error rate below 5%.

Statistical models and reporting

Dichotomous outcomes were compared between the two groups using a generalised linear model with binomial error distribution and a log-link to calculate an RR, and an identity-link to calculate a risk difference (RD). Both analyses were adjusted for the stratification variables (site, COPD, and haematological cancer). Due to a non-normal distribution of data for the outcomes 'percentage of days alive without life support', and 'percentage of days alive out of hospital' groups were compared using the non-parametric van Elteren test with adjustment for site only. The 'number of patients with one or more SAE' was compared similarly to the primary outcome. The Bayes factor (BF)[141] for the primary outcome was calculated using the freely available Bayes factor calculator (Bayes Factor Calculator, Copenhagen Trial Unit, Copenhagen, Denmark; ctu.dk/tools-and-links/bayes-factor-calculation).[142] The BF is the probability of obtaining the result from a trial (or meta-analysis) given the null-hypothesis is true (i.e. no difference) dived by the probability of obtaining the result of the trial given the alternative hypothesis is true (i.e. the a-priori postulated effects size). A low BF (<1) indicates that the obtained result is 1/BF times more likely to correspond to the alternative hypothesis than the null hypothesis. Conversely, a high BF (>1) indicates that the result is BF times more likely to correspond to the null hypothesis than the alternative hypothesis.[85]

A secondary analysis of the primary outcome, with adjustment for important prognostic baseline characteristics (age, type of ICU admission, sequential organ failure assessment (SOFA) score,[143] and absence or presence of metastatic cancer), in addition to the stratification variables was also conducted. This analysis was planned to be performed in a similar fashion as the primary analysis, but due to non-convergence in the pre-planned statistical model, this analysis was done using a logistic regression model (generalised linear model with binomial error distribution and a logit-link), and reported as an odds ratio (OR) with corresponding 95% CI.

Dichotomous variables were presented as numbers and percentages, whilst continuous variables were presented as means and standard deviations (SD) or medians and inter-quartile ranges (IQR), as appropriate.

3.1.9. Protocol adherence and sensitivity analyses

As protocol maintenance was imperative to ensuring proper separation of applied oxygenation targets, the ability to continuously monitor any protocol violations was paramount.

A major protocol violation (MPV) was defined as follows: both the highest and the lowest registered PaO_2 in one 12-hour interval (from 06:00 to 18:00 or from 18:00 to 06:00) were at least 1.0 kPa *above* the allocated PaO_2 target if both of the corresponding FiO_2 values were above 0.21 OR at least 1.0 kPa *below* the allocated PaO_2 target if both of the corresponding FiO_2 values were below 1.00. Thus, any MPV would correspond to the patient being off-target (either above or below) for at least one 12-hours interval, without maximum effort in regulating the FiO_2 being implemented. To support protocol adherence, e-mail notifications (to sponsor, co-ordinating investigators, and local site investigator) were automatically generated if any MPV occurred (as registered in the eCRF), and all notifications were systematically evaluated. If necessary, the responsible trial site would be contacted in order to investigate any safety concerns/issues in relation to the oxygenation target or the need for clarification of the trial protocol or additional support.

For sensitivity purposes of the mortality effect, four pre-planned per-protocol populations were defined,[123] being all patients except:

1. Those with an MPV deviating to the same side (above or below the allocated oxygenation target) in two or more consecutive 12-hours intervals
2. Those with one or more MPV
3. Those allocated to 12 kPa with MPVs deviating *below* the oxygenation target in two or more consecutive 12-hour intervals AND those allocated to 8 kPa with MPVs deviating *above* the oxygenation target in two or more consecutive 12-hour intervals
4. Those allocated to 12 kPa with MPVs deviating *above* the oxygenation target in two or more consecutive 12-hour intervals AND those allocated to 8 kPa with MPVs deviating *below* the oxygenation target in two or more consecutive 12-hour intervals

The conduct of sensitivity analyses of the HOT-ICU trial are currently ongoing, and the results from theses analyses are thus not presented in this thesis.

3.1.10. Trial site support

For supportive purposes a trial e-mail and a telephone hot-line (open 24/7/365 and staffed by the co-ordinating investigators [TLK or OLS] or trial sponsor [BSR]) was established. Trial sites were encouraged to make contact in case of any trial related questions. Contact information was available on the trial's website (cric.nu/hot-icu). Here, all relevant documents (e.g. protocols, newsletters, standard forms, patient information material, pocket cards, etc.) were also available, and the trial website was regularly updated. Newsletters were sent out on a monthly basis to all involved in the trial to inform on trial progress and important updates.

3.1.11. Data registration

HOT-ICU trial data was entered into an encrypted, web-based, password protected electronic case report from (eCRF). The system was supplied by the Copenhagen Trial Unit and used the clinical data management system OpenClinica© software (OpenClinica, LLC, Waltham, MA 02451, USA).

3.1.12. Data monitoring

Full external monitoring of consents and registered data at all trial sites was applied. This was done using a monitoring plan developed in collaboration with the Good Clinical Practice (GCP) unit at Aarhus and Aalborg Universities as according to the GCP standards.[144] Additional central monitoring was performed by the sponsor using eCRF data only. The trial was overseen by an independent data management and safety committee (DMSC). One pre-planned interim analysis of the trial was performed after inclusion of 50% of the patients (n = 1,462).

3.1.13. Ethical considerations

The HOT-ICU trial was approved by the Danish Health and Medicine Agency (Project ID AAUH-ICU-01, 2017-000632-34), the Danish Data Protection Agency (Project ID 2017-55), the Committee on Health Research Ethics in the North Denmark Region (Project ID N-20170015), and by all required authorities in all participating countries. All patients were enrolled in the trial after consent to participate had been obtained as according to national regulations.

3.1.14. Safety

By agreement with the Danish Medical Authorities, the occurrence three different SAEs during admission at participating ICUs was prospectively registered in the eCRF: new myocardial ischaemia, new ischaemic stroke, and new intestinal ischaemia. In the event of an SAE that was deemed either 'related' or 'possibly related' to the allocated oxygenation target an automatic e-mail notification was generated and sent to the sponsor, co-ordinating investigators, and site investigators at the specific site. This allowed for rapid evaluation of the clinical circumstances and handling of any safety concerns as required by law.

Based on registered data in the eCRF on the use of vasopressors/inotropes and plasma lactate levels, the occurrence of new episodes of shock was also evaluated and reported annually to the DMSC. Detailed definitions of SAEs are available elsewhere.[123,132]

3.2. Bayesian analysis of the HOT-ICU trial (Paper II)

3.2.1. Design

The protocol and statistical analysis plan for this study was published prior to inclusion of the last patient in the HOT-ICU trial.[96] All patients in the HOT-ICU intention-to-treat cohort, being all patients with data on 90-day all-cause mortality, were included.

The setup of the study was three-fold: to 1) probabilistically assess the overall mortality effect of a lower versus a higher oxygenation target; 2) estimate the probabilities for a range of clinically relevant treatment effects; and 3) based on pre-specified baseline variables, investigate the presence of heterogeneous treatment effects (HTE).

HTE was assessed according to the following baseline variables: 1) severity of illness as measured by the SOFA score; 2) severity of hypoxaemic respiratory failure as measured by the PaO_2/FiO_2 ratio; and severity of circulatory failure as measured by 3) vasopressor requirements represented by the highest dose of norepinephrine in the 24 hours prior to randomisation, and 4) latest plasma lactate concentration prior to randomisation. Rationale for choice of subgrouping schemes are elaborated in the protocol for this study and in Paper II (Appendix B).[96,133]

HTE was evaluated both according to subgroups of the outlined baseline variables, and on the continuous log-OR scale assessing interaction between the variable of interest and oxygenation target allocation (low versus high).

3.2.2. Outcome

For the secondary, Bayesian analysis of the HOT-ICU trial, only the primary outcome of the trial was considered: 90-day all-cause mortality.

3.2.3. Statistical analyses

Statistical software

For fitting the Bayesian statistical models we used Stan[145] version 2.26, accessed via the *brms* R package[146,147] version 2.15.0 and the *rstan* R package[148] version 2.21.2 using the open-source freely available software R version 4.0.4 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria). R Studio version 1.4.1106 was

used as interface for R (RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA).

Priors

We used three different priors for this study: for the primary analysis *weakly informative* priors, that encompassed all plausible effects sizes; for sensitivity purposes both *sceptic* priors (centred on no difference in outcomes and sceptical of large effect sizes), and *evidence-based* priors (informed by an updated random-effects meta-analysis of previously published RCTs) were used. Additional details on priors are available in the study protocol and in Paper II (Appendix B).[96,133]

Analysis of the primary outcome

For the analysis of the primary outcome, Bayesian logistic regression models with adjustment for the stratification variables of the HOT-ICU trial (site, COPD, and haematological cancer) were used. Results from this analysis were presented as median RRs, RDs, and ORs with corresponding 95% percentile-based CrIs. The full posterior distributions were presented visually. In addition, the probabilities of a range of effect sizes for the RR were calculated.

Subgroup-based HTE analyses

HTE was assesses according to the above-mentioned four sub-grouping schemes, each with five quintile-based subgroups, and investigated using hierarchical Bayesian logistic regression models adjusted for the stratification variables. The RRs, RDs, and ORs were calculated and summarised as outlined above. Additional adjustment for the type of oxygen supplementation system used, (open or closed), was performed when considering the baseline PaO_2/FiO_2 ratio. Posterior probability distributions were presented similarly to the analysis of the primary outcome.

Continuous HTE analyses

The potential interaction of the allocation to the lower oxygenation target with the four baseline characteristics of interest for 90-day all-cause mortality on the continuous log-OR scale was investigated using Bayesian logistic regression models. Treatment allocation was included, and all models were adjusted for the stratification variables as well as adjustment for type of oxygen supplementation system when assessing PaO_2/FiO_2 ratio. ORs for interactions, and probabilities for interaction-ORs <1 and >1 were also presented. An interaction-OR less than one indicated negative interaction, i.e. decreased risk of death with the lower oxygenation target and increasing levels of the selected baseline parameter, as compared with the higher oxygenation target. Conversely, an interaction-OR more than one indicated positive interaction, i.e. increased risk of death with the lower oxygenation target and increasing levels of the selected baseline parameter, as compared with the higher oxygenation target.

The predicted probabilities, dependent on the variables in focus, were presented visually as conditional effects plots.

Handling of missing data

If ≥5% data were missing for any variable in any analysis multiple imputation was planned. Otherwise a complete case analysis would be performed.

3.3. Updated Cochrane review (Paper III)

This paper represents an update of a previously published Cochrane review,[30] which in turn is based on a detailed, pre-defined review protocol.[149] The methodology presented in this thesis is based on the same methodology as the first review, and will thus to great extents be similar, though small changes were made. These are summarised in the section 'Differences between protocol and review' in Paper III (Appendix C).[134]

3.3.1. Eligibility criteria

Study selection

RCTs, irrespective of reported outcomes, publication status, publication date, and language were included for consideration. Unpublished trials would only be included if methodological descriptions and trial data could be obtained by direct contact with trial authors or in written form. Randomised cross-over trials, and quasi-randomised trials were excluded.

Participants

Adults ≥18 years admitted to the ICU prior to randomisation were included.

Interventions

Trials with a clear differentiation of participants randomised to a lower or a higher oxygenation strategy were included. Participants mechanically ventilated (including IMV, NIV, CPAP) or non-mechanically ventilated (all open oxygen supplementation systems, including high-flow systems) were eligible.

The control (or 'comparator') group was defined as adults receiving a lower ('conservative') oxygenation strategy. This could be achieved via any oxygen supplementation device. The aim of the group would be minimising exposure to hyperoxia in the lungs. This could be achieved by exposure of the participant to low levels of FiO_2, or low targets of PaO_2/SaO_2/SpO_2.

The intervention (or 'experimental') group was defined as adults receiving a higher ('liberal') oxygenation strategy. As in the control group, this could be achieved by any oxygenation device. The aim in this group would be ensuring adequate oxygenation through exposure to hyperoxia in the lungs. This could be achieved by either by high levels of FiO_2, or high targets of PaO_2/SaO_2/SpO_2.

Trials or groups randomised to hypoxaemia (i.e. an FiO_2 <0.21, SaO_2/SpO_2 < 80%, or PaO_2 <6 kPa), or hyperbaric oxygen were excluded.

3.3.2. Outcomes

In Cochrane reviews, a maximum of seven pre-defined, clinically important outcomes can be specified to inform the GRADE assessment (Grading of Recommendations, Assessment, Development and Evaluations),[126,129] and for this review three co-primary and four secondary outcomes were pre-defined.[149] Table 1 summarises the pre-specified outcomes. All outcomes were reported at maximum follow-up as defined by trialists.

Table 1. Pre-specified outcomes in the Cochrane review update

Level	Outcome
Co-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)**
	Myocardial infarction
	Stroke
	Sepsis

**Serious adverse events were defined as "any untoward medical occurrence that resulted in death, was life-threatening, required hospitalisation or prolongation of existing hospitalisation, resulted in persistent or significant disability, or jeopardised the participant".*[149]
***Defined as either pneumonia, acute respiratory distress syndrome (ARDS), or pulmonary fibrosis, or as defined by trialist.*

3.3.3. Search methods

The following databases were searched: Cochrane Central Register of Controlled Trials; MEDLINE; Embase; Science Citation Index; BIOSIS Previews; and Latin American and Caribbean Health Science Information database. Cumulative Index to Nursing and Allied Health Literature was searched for the primary edition of this review,[149] but not for this updated version due to restrictions of access.

Ongoing and unpublished trials were searched in the following trial registers: US National Institutes of Health Ongoing Trials Register; World Health Organization International Clinical Trials Registry Platform; EU Clinical Trials Register; and Australian New Zealand Clinical Trials Registry.

Finally, reference lists of included trial reports, relevant reviews, or papers of randomised and non-randomised trials, and editorials were manually screened for potentially relevant trials.

For details on the search strategy applied please refer to Paper III (Appendix C).[134]

3.3.4. Data collection and analysis

Two authors screened each title and abstract of all reports identified by the searches, and subsequently potentially relevant full texts reports were obtained and assessed for inclusion. The task was performed independently and in pair. Any disagreements were resolved by consensus or by consulting another author. Table 2 summarises the data extraction form. Corresponding authors of included trials were contacted in case of a need for clarification of issues relating to data reporting or if additional study details were needed.

Risk of bias

All outcomes were assessed for risk of bias by two independent authors, as according to the Cochrane Handbook for Systematic Reviews of Interventions,[126] using the 'Risk of Bias 2 tool' (RoB-2).[128] Any disagreements were resolved by discussion or consultation with a third author.

Risk of bias was assessed in the following domains: 1) randomisation process; 2) deviations from the intended interventions; 3) missing outcome data; 4) measurement of the outcome; and 5) selection of the reported result. Each domain was adjudicated as being at 'low risk of bias', 'some concerns', or 'high risk of bias'. Included trials judged to be at low risk of bias if all domains were classified as being at overall at low risk of bias. Those with one domain assessed as being of some concerns, but no domain judged as being at high risk of bias, where designated as being overall at some concerns of risk of bias. Trials were classified as being at overall high risk of bias if one or more domains were judged as being at high risk of bias. However, if any trial was judged as being of some concerns in multiple domains, and the assessors judged that the multiple concerns amounted to a serious risk of bias, the trial was judged as being at overall high risk of bias.

Table 2. Trial data extraction

Domain	Information
Trial	Country
	Duration of the trial
	Date of publication
	Type of trial
Participants	Numbers randomised
	Numbers analysed
	Numbers lost to follow-up or withdrawn
	Type of population
	Mean or median age
	Sex
	Inclusion criteria
	Exclusion criteria
Interventions	Intervention
	Comparator
	Concomitant interventions
Outcomes	Pre-defined co-primary and secondary outcomes

Meta-analysis

All meta-analyses were performed using the statistical software Review Manager Web (RevMan Web, version 3.6.0, The Cochrane Collaboration, 28 June 2021, available at revman.cochrane.org) and the TSA software version 0.9 (The Trial Sequential Analysis, Copenhagen Trial Unit, Copenhagen, Denmark. The software is freely available at ctu.dk/tsa). Intervention effects were assessed using both fixed-effect and random-effects models, and the most conservative estimate was reported.[140] Results of the meta-analyses were presented visually by forest plots.

Statistical significance

As the higher oxygenation strategy was considered the experimental group ('intervention'), RRs <1 would indicate benefit of a higher oxygenation strategy, whereas RRs >1 would indicate benefit of the lower oxygenation strategy ('comparator'). RRs with 95% CIs were calculated for dichotomous outcomes, and mean differences with 95% CIs were calculated for continuous outcomes. Multiplicity adjustment of the significance levels was performed for all outcomes, as suggested by Jakobsen et al.[140]

Trial sequential analysis

It is generally recommended, and specifically by Cochrane, to update systematic reviews and meta-analyses at regular intervals or when new trials are published.[126] However, the risk of a type-I error (i.e. false positive: claiming a finding is statistically

significant when in fact it is not and only due to chance) is increased whenever new trials are added to an analysis.[150] To account for this, it is possible to apply a Trial Sequential Analysis (TSA) wherein addition of each trial is considered a separate interim analysis, the multiplicity is accounted for, and thus conserving the family-wise-significance level.[151,152] The required information size (i.e. required number of participants) to confirm or reject an a priori defined effect size can also be calculated. This calculation takes into account the proportion of events in the control group, the proposed effect size, and the statistical variance within the meta-analysis.[153] The TSA can also be used to test for significance whenever a new trial is added to the analysis, and does so by constructing 'trial sequential monitoring boundaries'. This enables statistical inference based on cumulative meta-analyses that have not yet reached their required information size.[150,154,155] If the trial sequential monitoring borders are crossed before the required information size is reached, it is possible to establish firm evidence for benefit or harm. This will render further investigations redundant. However, if the boundaries are not crossed it may be prudent to continue with additional trials before certainty of effect can be established. If the trial sequential monitoring boundaries for futility are crossed the TSA may be used to evaluate the lack of a postulated effects size. However, effects sizes smaller than the tested may still be plausible, and if clinically relevant may support the conduct of additional trials.

All co-primary and secondary outcomes were analysed using TSA, and TSA was also used to estimate the required information sizes for each outcome. Further, TSA CIs were calculated, and analyses were presented graphically by means of TSA-plots.

Bayes factor

The BF was calculated for all outcomes similarly as described previously (see section 3.1.8).

Heterogeneity

Signs of statistical heterogeneity were assessed by visual inspection of the forest plots, by a significance set at $p < 0.10$ from the Chi^2 test, and by the I^2 statistic.[156]

Clinical diversity and subgroups

Potential clinical diversity (i.e. differences in trial design, included participants, outcome measurements and definitions, etc.) was investigated by means of the subgroup analyses summarised in Table 3.

Table 3. Subgroups

Grouping scheme	Subgroups
Risk of bias	Overall 'low risk of bias'
	Overall 'some concerns'
	Overall 'high risk of bias'
Oxygen intervention	Oxygenation target measured using either PaO_2 or SaO_2 or SpO_2 (as defined by trialists)
	Oxygen level defined by FiO_2 (as defined and set by trialists)
	Difference between groups (as defined by trialist)
FiO_2 or oxygenation target in higher group	Low targets
	High targets
FiO_2 or oxygenation target in lower group	Low targets
	High targets
ICU population	Medical
	Surgical
	Mixed
	Any respiratory failure
	Any cerebral disease
	Any heart disease
	Any trauma
	COPD
Oxygen delivery system	Invasive mechanical ventilation
	Non-invasive oxygen administration
	Mixed oxygen delivery system

Sensitivity analyses

The following sensitivity analyses were conducted: 1) trials at overall judged to be at low risk of bias only; 2) 'best-worst-case' scenario assuming participants lost-to follow-up in the higher group did not have an event, and participant lost-to follow-up in the lower group had an event; and 3) 'worst-best-case' scenario assuming participants lost-to follow-up in the higher group did had an event, and participant lost-to follow-up in the lower group did not have an event. Two post-hoc defined sensitivity analyses of the occurrence of SAEs and lung injuries were also conducted: 1) the highest reporting proportion of an event, and 2) estimated cumulated number of events.

Assessment of the certainty of the evidence

For each outcome, two independent authors assessed the certainty of evidence by using the GRADEpro GDT tool which is integrated in the RevMan Web tool.[157] Outcomes were presented first for trials judged to be at overall low risk of bias only, and for all included trials secondly.

4. RESULTS

4.1. The HOT-ICU trial (Paper I)

4.1.1. Recruitment

The first patient was recruited on June 20, 2017 and recruitment was completed on August 3, 2020. See Figure 1 for recruitment rates.

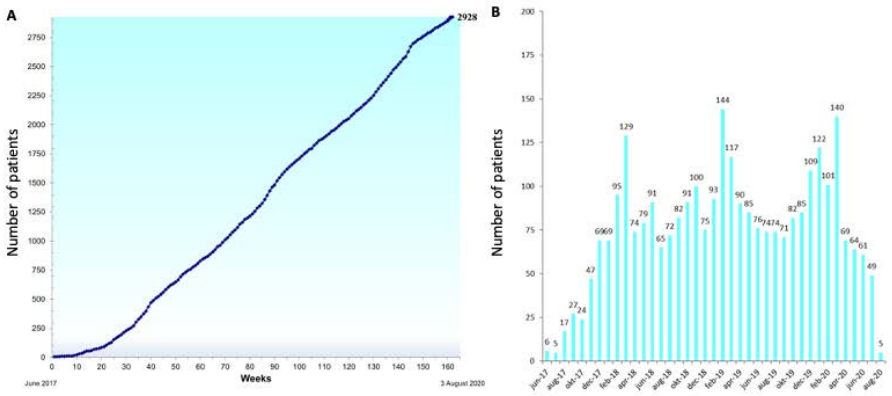


Figure 1. Inclusion rate of the HOT-ICU trial. ***A****: Number of patients recruited per week.* ***B****: number of patients recruited by month.*

Recruitment was slow in the beginning, but increased as new sites were initiated. There was a striking seasonal variation in recruitment rates, with higher rates during winter and early spring, and lower rates during summer. With the advent of the SARS-CoV-2 pandemic (COVID-19), taking its effect on HOT-ICU trial sites by late February/early March 2020, an initial dramatic increase in the recruitment rate was observed. This was however subsequently reduced as many trial sites halted their participation due to lack of staff.

A total of 4,192 patients were screened, of which 2,928 patients were included from 35 ICUs in 7 countries (Denmark, Switzerland, Finland, the Netherlands, Norway, the United Kingdom, and Iceland). However, one site only managed to screen patients. See Figure 2 for the number of patients included at each site.

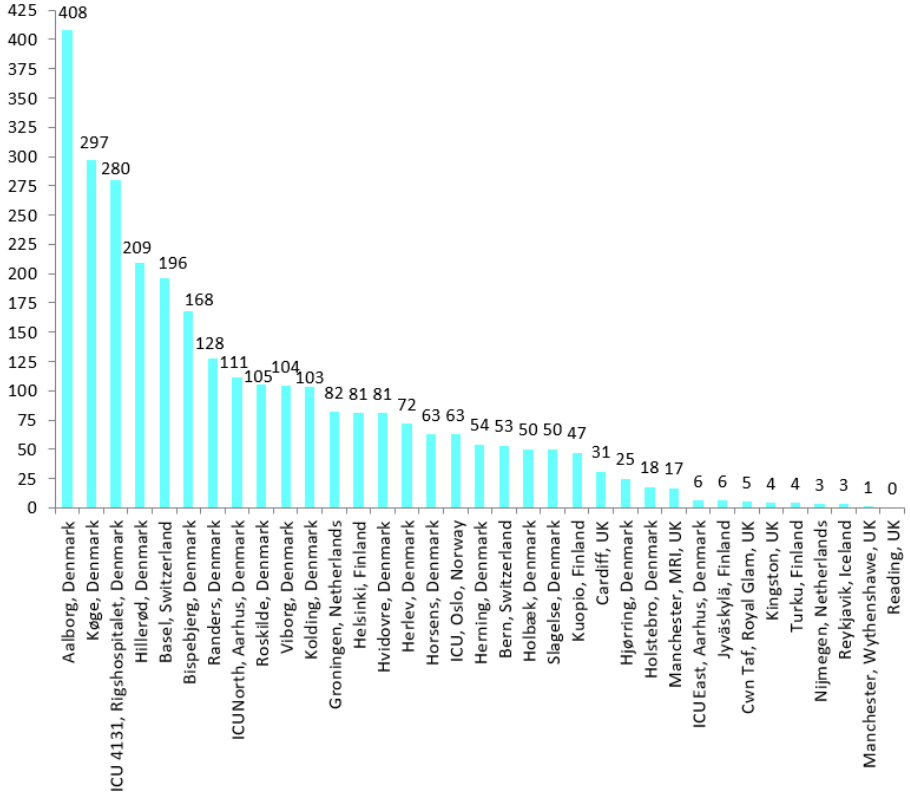


Figure 2. Number of patients recruited at each site in the HOT-ICU trial.

4.1.2. Trial population

Due to minimal loss to follow-up or withdrawal of consent/consent not obtainable, 2,888 patients were included in the intention-to-treat analysis (98.6%). No imputation for missing data was performed as the level of missing data was less than 5% for any variable included in any analysis. The two groups were comparable at baseline, except for incidence of cardiac arrest prior to randomisation. The CONSORT diagram and baseline characteristics are available in Paper I (Appendix A).[132]

4.1.3. Oxygenation parameters

The highest and lowest PaO_2 were registered for pre-defined 12-hour intervals with concomitant measurements of FiO_2 and SaO_2. A clear separation in all oxygenation parameters (PaO_2, FiO_2 and SaO_2) was achieved when comparing the two groups during the entire 90-day intervention period.

Arterial oxygen partial pressure

Figure 3 represents the daily median patient-mean PaO_2 registrations until day 90 after randomisation. Daily patient-means were calculated from the registered 12-hour highest and lowest PaO_2 measurements. Median PaO_2 was 12.4 kPa (IQR: 11.6-13.2 kPa) in the higher group and 9.4 kPa (IQR: 8.9-10.2 kPa) in the lower group for the entire intervention period. For the entire 90-day intervention period, only 2.3% of patients in the lower group had a median PaO_2 below 8 kPa, but 35.3% of patients in the higher group had a median PaO_2 below 12 kPa.

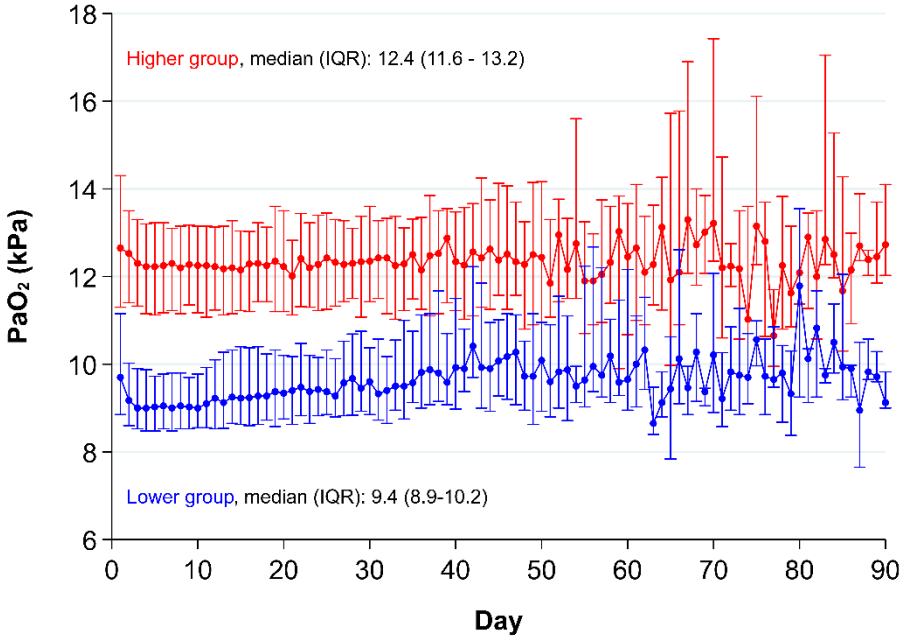


Figure 3. Median values of daily patient-means of partial pressure of arterial oxygen (PaO_2) stratified according to oxygenation target allocation for the 90-day intervention period. Daily patient-means were calculated from the registered 12-hour highest and lowest values. Bars represents interquartile ranges (IQR). The y-axis has been changed from mmHg to kPa. Adapted from Paper I and published with permission from the journal.[132]

Figure 4 represents the distribution of PaO_2 registrations. In total, 20.9% of registered PaO_2 measurements in the lower group were below 8 kPa, whilst 53.8% in the higher group were below 12 kPa. In the lower group 86.8% of patients had one or more registered PaO_2 measurement below 8 kPa, and in the higher group 99.4% of patients had one or more registered PaO_2 measurement below 12 kPa. There was a markedly right skewed distribution in PaO_2 values in the lower group, whereas the values were more normally distributed in the higher group.

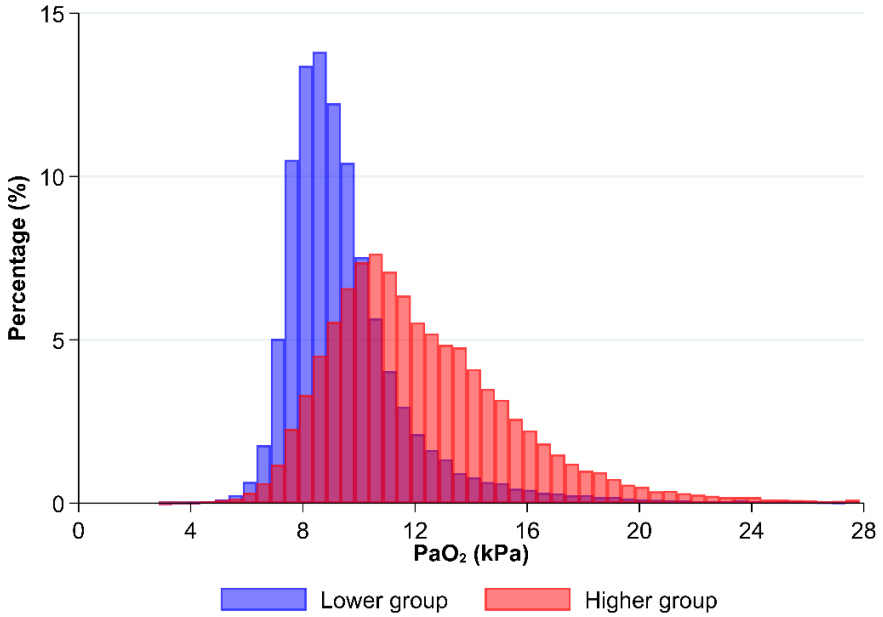


Figure 4. Histogram of registered 12-hour highest and lowest partial pressures of arterial oxygen (PaO_2) in kPa. Bars indicate the percentage of registrations in each group. Data are censored at PaO_2 ≤28 kPa and include 103,836 of 104,567 registrations (99.3%).

Fraction of inspired oxygen

Figure 5 represents the daily median patient-mean FiO_2 settings until day 90 after randomisation. Median FiO_2 was 0.56 (IQR: 0.46-0.71) in the higher group and 0.43 (IQR: 0.34-0.54) in the lower group for the entire intervention period (Figure 5). Patient-means are calculated from the FiO_2 settings registered concomitantly to the 12-hour highest and lowest PaO_2 measurements. At randomisation, the FiO_2 in both groups was 0.70, but as is evident from Figure 5, this was rapidly reduced, and separation of the two groups was already seen from day 1.

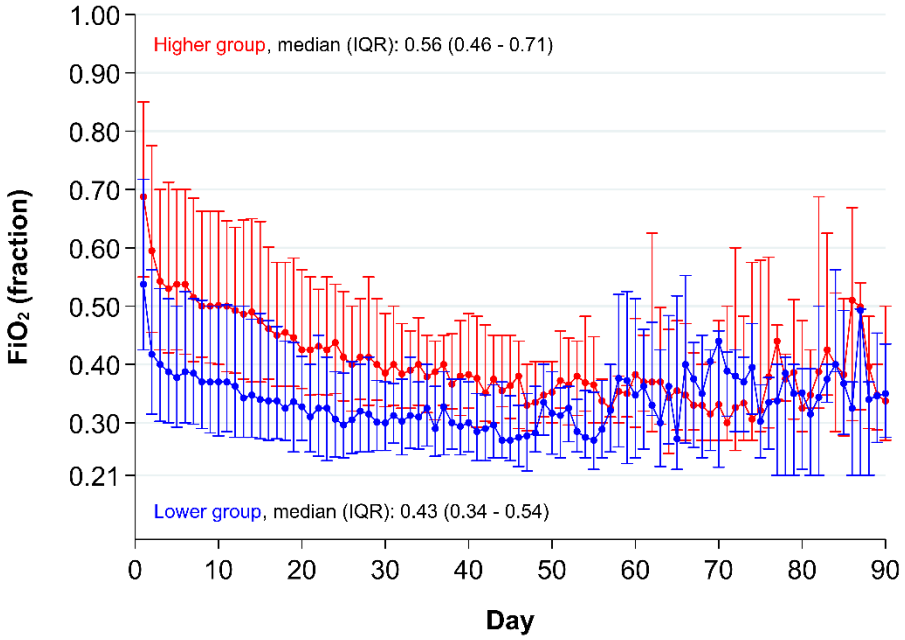


Figure 5. Median values of daily patient-means of fractions of inspired oxygen (FiO_2) stratified according to oxygenation target allocation for the 90-day intervention period. Daily patient-means were calculated from the registered values corresponding to the 12-hour highest and lowest partial pressure of arterial oxygen. Bars represents interquartile ranges (IQR). Published with permission from the journal.[132]

The FiO_2 settings used by clinicians to achieve the two oxygenation targets were clearly different in the two groups, with a tendency for higher registered fractions in the higher group as compared to the lower group (Figure 6). Despite this, most registered PaO_2 values had concomitant fractions below 1.00; in the lower group, 96.3% of registered FiO_2 were less than 1.00. This was the case for 92.3% of FiO_2 in the higher group (Figure 6). A total of 11.3% of all recorded PaO_2 values in the lower group were at an FiO_2 of 0.21, whilst this was the case for 1.8% in the higher group.

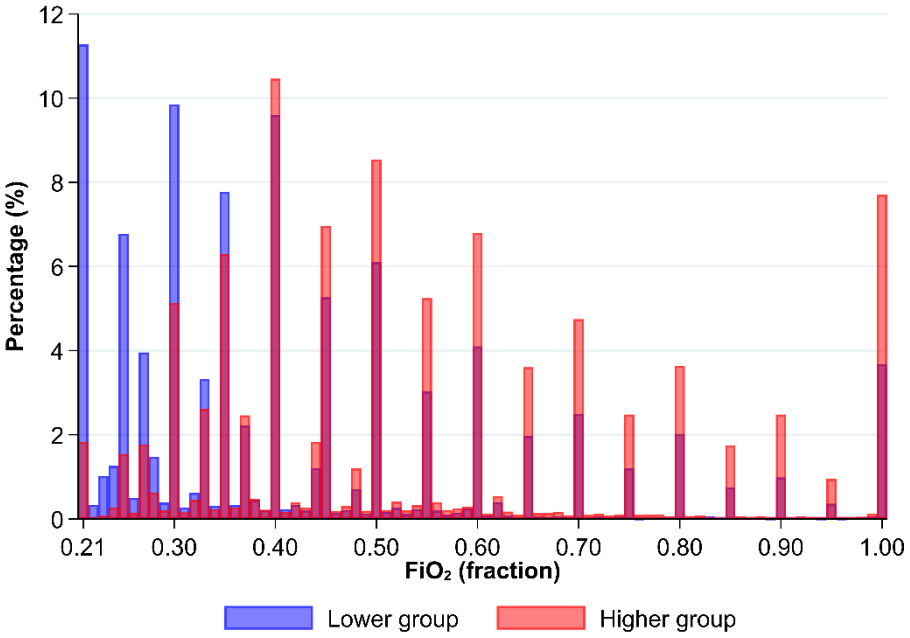


Figure 6. Histogram of all registered 12-hour highest and lowest fractions of inspired oxygen (FiO_2) during the 90-day intervention period. Bars indicate the percentage of registrations in each group. Based on 104,430 registrations.

Arterial oxygen saturation

Figure 7 represents the daily median patient-mean SaO_2 registrations until day 90 after randomisation. Median SaO_2 was 96% (IQR: 95-97%) in the higher group and 93% (IQR: 92-94%) in the lower group for the entire intervention period (Figure 7). Daily patient-means are calculated from the SaO_2 measurements registered concomitantly to the 12-hour highest and lowest PaO_2 measurements.

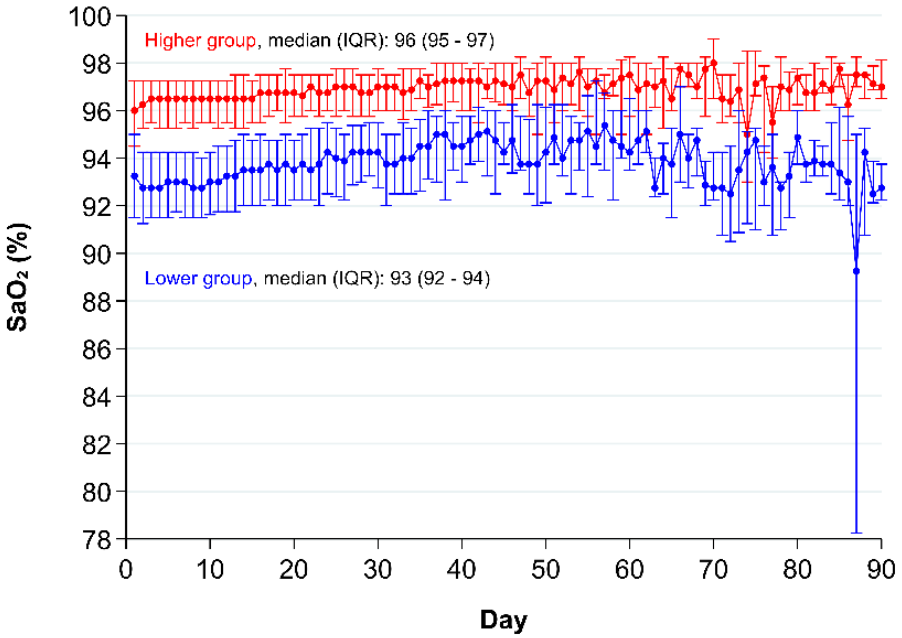


Figure 7. Median values of daily patient-means of arterial oxygen saturation (SaO2) stratified according to oxygenation target allocation for the 90-day intervention period. Daily patient-means were calculated from the registered values corresponding to the 12-hour highest and lowest partial pressure of arterial oxygen. Bars represents interquartile ranges (IQR). Published with permission from the journal.[132]

A near-normal distribution of SaO_2 measurements in the lower group was observed, but the distribution was noticeably left skewed in the higher group (Figure 8). In the lower group, 0.8% of SaO_2 registrations were equal to 100%. This was the case for 2.7% of SaO_2 registrations in the higher group. However, both groups where naturally limited to a maximum of 100%.

In the lower and higher groups, 0.06% and 0.04% of SaO_2 registrations were less than 80%, respectively. However, 14.8% of patients in the lower group had one or more SaO_2 registration less than 80%. This was the case for 9.8% of patients in the higher group.

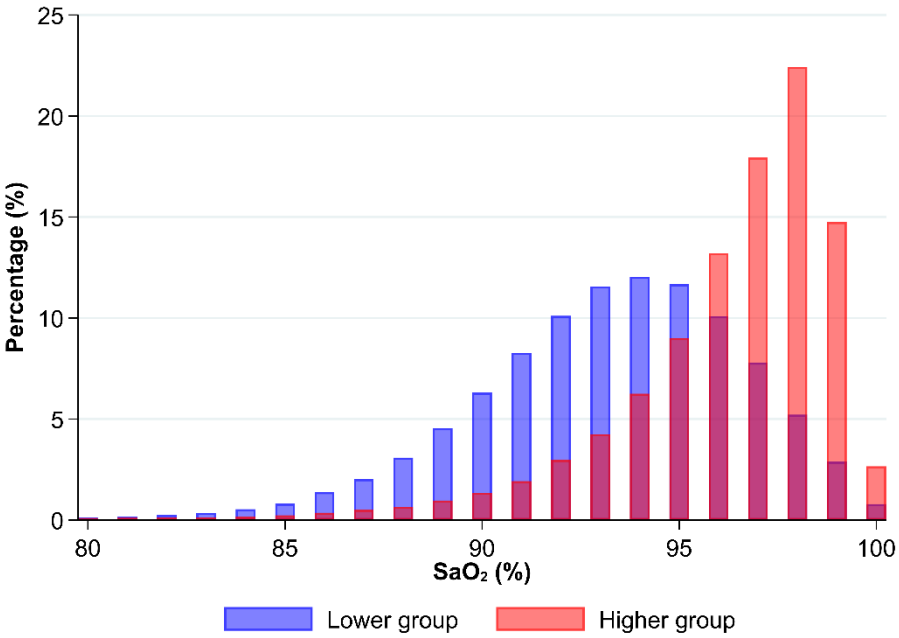


Figure 8. Histogram of registered 12-hour highest and lowest arterial oxygen saturation (SaO_2) during the 90-day period. Bars indicate the percentage of registrations in each group. Data are censored at SaO_2 ≥80% and include 97,168 of 97,734 registrations (99.4%).

Arterial partial pressure of oxygen versus arterial oxygen saturation

The relationship between all PaO_2 registrations and their concomitant SaO_2 measurements is displayed in Figure 9 (data is not stratified for group allocation). There was a substantial spread of data as one measurement of PaO_2 could reflect several different SaO_2 values and vice versa. For the sake of clarity, data from the lowest region of the PaO_2 spectrum (0 to 14 kPa) have been magnified in the inserted section of Figure 9.

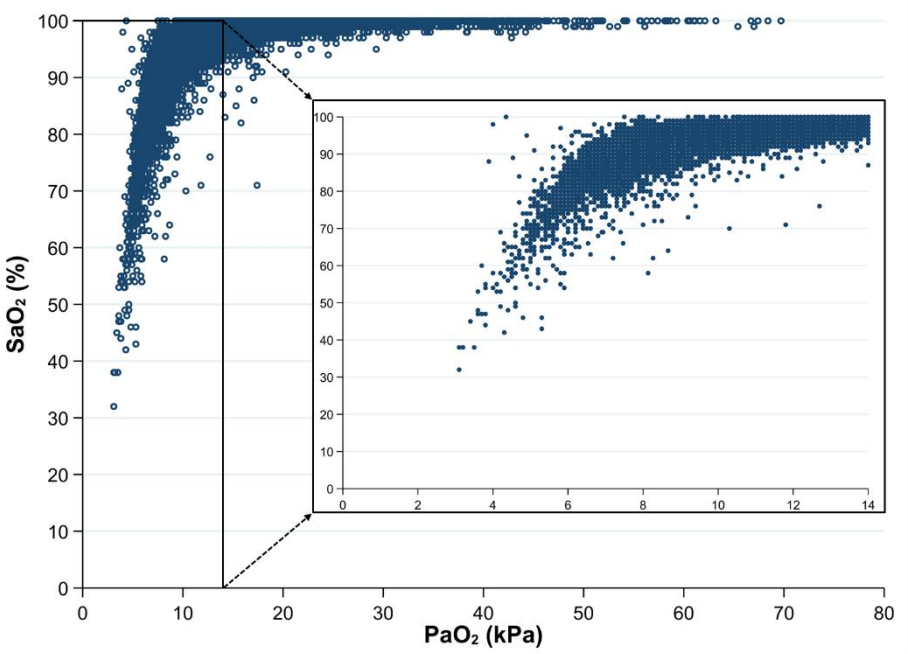


Figure 9. Relationship between all registrations of arterial partial pressure of oxygen (PaO_2) and concomitant arterial oxygen saturation (SaO_2). SaO_2 measurements were not performed at one site, thus data from 191 patients were missing. The graph is based on 97,728 point estimates. Data are censored at ≤14 kPa in the framed sub-graph (82,184 point estimates = 84.1%).

PaO_2/FiO_2 ratio

Figure 10 represent the daily patient-mean PaO_2/FiO_2 ratios until day 90 after randomisation. At baseline, the two groups had similar PaO_2/FiO_2 ratios: 15.8 kPa (IQR 11.8-21.0 kPa) in the lower group and 15.7 kPa (IQR 12.0-20.5 kPa) in the higher group (Table 4). The vast majority of patients in both groups had ratios equivalent to moderate or severe ARDS at baseline (Table 4).

Table 4. Baseline PaO_2/FiO_2 ratios in the HOT-ICU trial

Baseline PaO_2/FiO_2 ratio	Lower group (%) (n = 1,448)	Higher group (%) (n = 1,450)
Ratio < 13.3 kPa	33.9%	34.5%
13.3 kPa ≤ ratio < 26.7 kPa	53.3%	54.8%
26.7 kPa ≤ ratio < 40 kPa	9.7%	7.3%
Ratio ≥ 40 kPa	3.0%	3.4%

Baseline PaO_2/FiO_2 ratios subdivided according to categories of the Berlin definition of ARDS.[41] *Presented as percentage of patients with available PaO_2/FiO_2 ratios at baseline.*

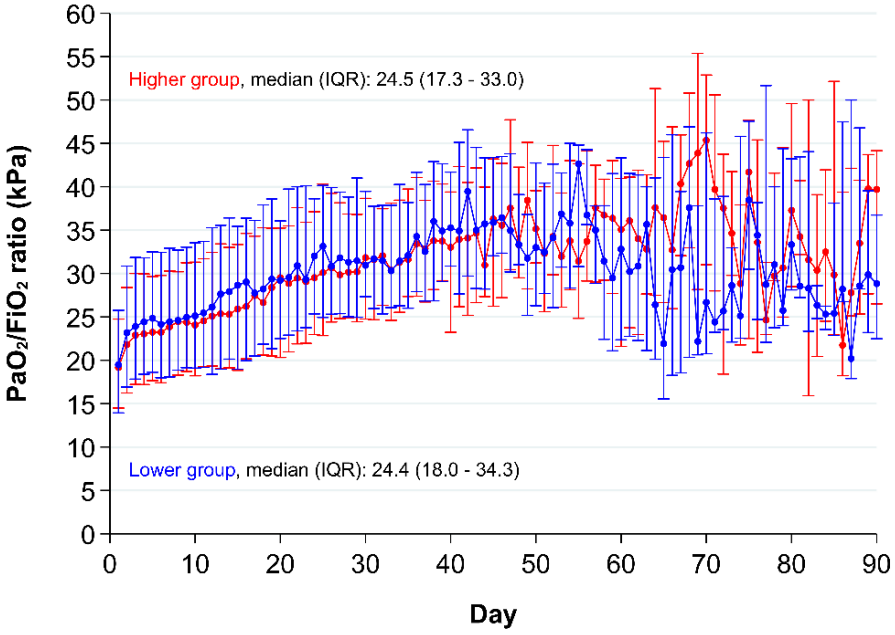


Figure 10. Median patient-mean partial pressure of oxygen to fraction of inspired oxygen (PaO_2/FiO_2) ratios during the 90-day period. Bars represent interquartile ranges.

A large proportion of patients in both groups had mean PaO_2/FiO_2 ratios below 40 kPa, corresponding to the upper cut-off for ARDS,[41] throughout the trial (Table 5).

Table 5. Patient-mean PaO_2/FiO_2 ratios for the entire intervention period

Patient-mean PaO_2/FiO_2 ratio	Lower group (%) (n = 1,444)	Higher group (%) (n = 1,456)
Ratio < 13.3 kPa	9.1%	7.8%
13.3 kPa ≤ Ratio < 26.7 kPa	49.8%	57.2%
26.7 kPa ≤ Ratio < 40 kPa	35.1%	30.5%
Ratio ≥ 40 kPa	6.0%	4.5%

Patient-mean PaO_2/FiO_2 ratios for the entire HOT-ICU intervention period, subdivided according to categories of the Berlin definition of ARDS.[41] Presented as percentages of patients with available PaO_2/FiO_2 ratios during the intervention period.

We observed a remarkable overlap in point estimates of PaO_2/FiO_2 ratios during the intervention period of the two groups (Figure 11).

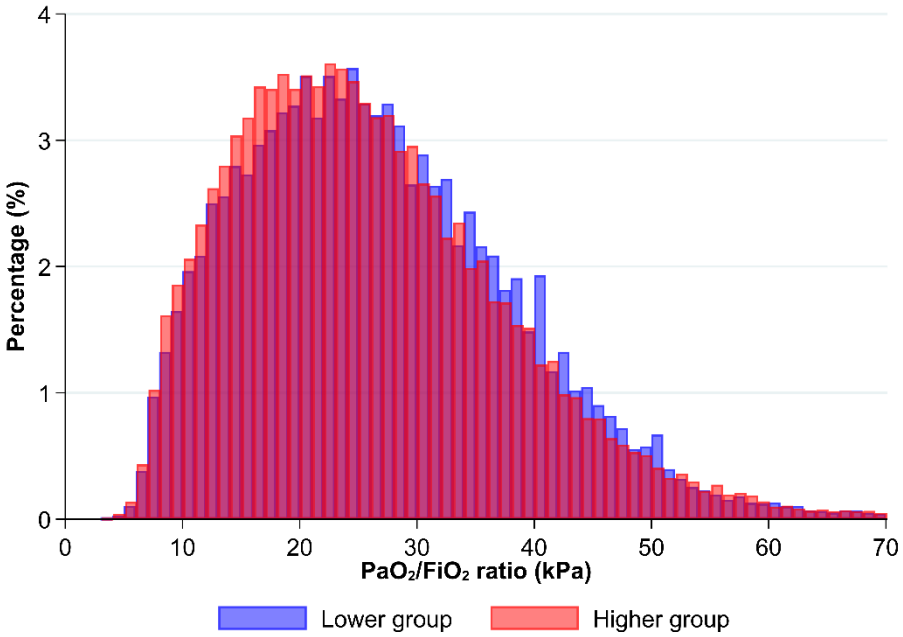


Figure 11. Histogram of point estimates of partial pressure of oxygen (PaO_2) to fraction of inspired oxygen (FiO_2) ratios during the 90-day period. Data are censored at PaO_2/FiO_2 ratio ≤70 kPa and include 103,796 of 104,425 point estimates (99.4%).

4.1.4. ICU-treatment

Similar use of mechanical ventilation (IMV or NIV), prone positioning, inhalation of vasodilators, ECMO, infusion of vasopressors or inotropes, renal replacement therapy, and red blood cell transfusions was seen in the two groups. Additionally, data on ventilation parameters obtained daily at 08:00 (peak inspiratory pressure, peak end-expiratory pressure, or tidal volume for patients undergoing IMV and end-expiratory pressure for patients undergoing NIV) were similar. The frequency of arterial blood sampling was also similar; with a mean of 6 per patient per day in both groups, with the highest rates at the beginning of the trial and with a steady reduction in both groups towards day 90 (Figure 12). Additional information on ICU treatment is available in Paper I (Appendix A).[132]

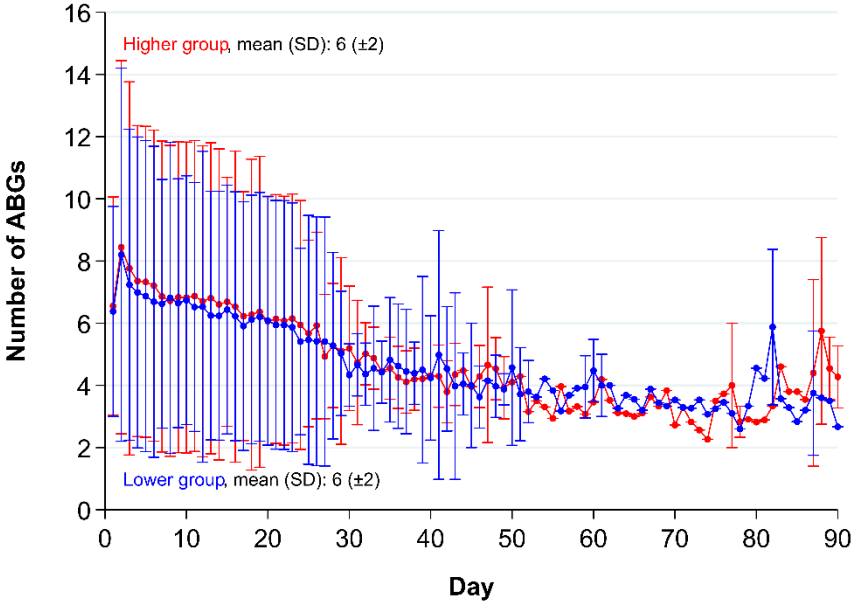


Figure 12. Mean number of arterial blood gas samples (ABGs) per day. Bars represent standard deviations (SD).

As is evident from Table 6, not all patients admitted to a participating ICU and providing data in general had recorded information on oxygenation parameters as measured by arterial blood gas sampling. The pattern over time was similar in the two groups, and with a relative increase over time (Figure 13). There was a seemingly exponential, but equal decline in the number of patients providing data on oxygenation (by means of 1 or more ABG) and number of patients overall during the intervention period in both groups (Table 6).

Table 6. Number of patients in the HOT-ICU trial providing data during trial conduct

	Lower group		Higher group	
Day	**Patients providing data**	**≥1 ABG**	**Patients providing data**	**≥1 ABG**
1	1453	1432	1457	1442
10	435	410	524	495
20	189	169	227	209
30	102	79	125	104
40	50	42	64	46
50	28	24	38	28
60	21	16	26	15
70	15	8	18	12
80	9	6	11	6
90	6	4	11	9

Number of patients admitted to a participating ICU and providing data for the trial, and number of patients reporting data on one or more arterial blood gas sample (ABG) conducted, stratified by treatment allocation. Listed by day after randomisation.

By day 6-7, roughly 50% of randomised patients remained in a participating ICU, and by day 22 this had dropped to around 10%.

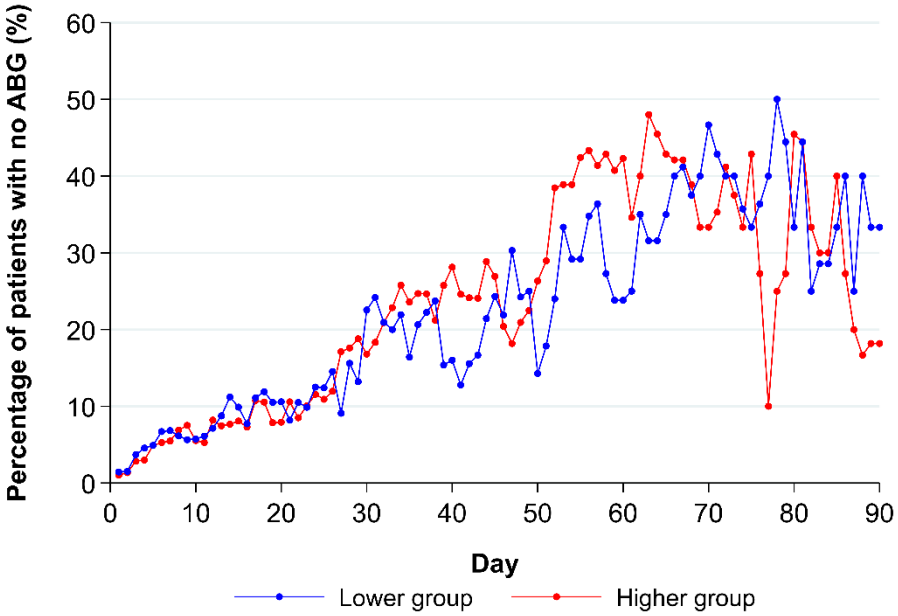


Figure 13. Percentage of patients admitted to a participating ICU but with analysis of an arterial blood gas sample (ABG) conducted on the given day.

4.1.5. Outcomes

Primary outcome

By 90 days post-randomisation, 613 of 1,447 patients (42.4%) in the higher-oxygenation group and 618 of 1,441 patients (42.9%) in the lower-oxygenation group had died, yielding an adjusted risk ratio (RR) of 1.02 (95% CI: 0.94-1.11; p = 0.64). Conducting additional adjustment for important baseline characteristics resulted in an OR of 1.06 (95% CI: 0.90-1.24; p = 0.50). Bayes factor for the primary outcome was markedly larger than 1, thus supportive of the null-hypothesis.

Secondary outcomes

No significant differences in secondary outcomes were found. See Paper I (Appendix A) for additional information.[132]

4.1.6. Subgroup analyses

A number of pre-specified and post-hoc defined subgroup analyses on 90-day all-cause mortality, based on selected baseline characteristics were conducted.[123,124] The results are displayed in Table 7. No significant interaction with oxygenation target allocation on 90-day all-cause mortality was found for any of the listed subgroups analyses.

In a post-hoc study of the 110 patients testing positive for SARS-CoV-2 (during or leading to index ICU admission) randomised in the HOT-ICU trial, all pre-defined outcomes were also investigated. No significant difference between the two groups was found (adjusted RR for 90-day all-cause mortality 0.87, 95% CI 0.58-1.38, p = 0.51). Further details are available elsewhere.[158]

Table 7. Subgroup analyses in the HOT-ICU trial

Characteristic, no. of events/no. of patients (%)	Lower group	Higher group	RR (95% CI)
Shock at baseline			
Yes	244/427 (57.1)	219/412 (53.2)	1.07 (0.95-1.21)
No	374/1014 (36.9)	394/1035 (38.1)	0.97 (0.87-1.08)
Invasive mechanical ventilation at baseline			
Yes	380/826 (46.0)	378/863 (43.8)	1.07 (0.97-1.19)
No	238/615 (38.7)	235/584 (40.2)	0.95 (0.83-1.09)
COPD			
Yes	122/277 (44.0)	132/285 (46.3)	0.98 (0.82-1.17)
No	496/1164 (42.6)	481/1162 (41.4)	1.03 (0.94-1.13)
Traumatic brain injury			
Yes	3/8 (37.5)	2/14 (14.3)	2.63 (0.55-12.54)
No	615/1433 (42.9)	611/1433 (42.6)	1.01 (0.93-1.10)
Cardiac arrest			
Yes	96/147 (65.3)	111/185 (60.0)	1.09 (0.92-1.28)
No	522/1294 (40.3)	502/1262 (39.8)	1.03 (0.94-1.13)
Type of admission			
Medical	540/1238 (43.6)	536/1233 (43.5)	1.02 (0.93-1.11)
Elective surgical	6/18 (33.3)	4/21 (19.1)	1.68 (0.56-5.06)
Acute surgical	72/185 (38.2)	73/193 (37.8)	1.13 (0.88-1.45)

Subgroup analyses according to baseline characteristics. COPD denotes chronic obstructive pulmonary disease, CI confidence interval, RR relative risk. An RR >1 favours the higher target whilst an RR <1 favours the lower target. Published with permission from the journal.[132]

4.2. Bayesian analysis of the HOT-ICU trial (Paper II)

4.2.1. Patient population

All 2,888 patients in the HOT-ICU intention-to-treat cohort were included in this study (98.6% of randomised patients). Baseline characteristics are presented in the supplement for Paper II (Appendix B).[133] No imputation for missing data was performed as the level of missingness for any parameter in any analyses was less than 5%.

4.2.2. Bayesian analysis of mortality

The RR for all-cause mortality at 90 days post-randomisation, adjusted for the stratification variables (site, COPD, and haematological malignancy) was 1.02 (95% CrI: 0.93-1.11). There was a 63.5% probability of an RR >1.00 (i.e. favouring the higher oxygenation target). The probability of an RR <0.80, corresponding to the a-priori hypothesised intervention-effect,[123] was less than 0.01%. Additional details on mortality effects sizes are provided in Paper II (Appendix B).[133]

4.2.3. Subgroup based HTE analyses

A potential benefit of the higher oxygenation target for increasing dose of continuously infused norepinephrine at baseline was suggested. RR in the lowest dosage group (dose = 0.0 mM) was 0.99 (95% CrI: 0.87-1.11), increasing to 1.08 (95% CrI 0.95-1.33) in the highest dosage group (dose: 0.40-2.40 µg/kg/min.).[133] No such potential dose-response relationship was suggested in any of the other subgrouping schemes.

Figure 14 illustrates the posterior probability distributions for the RR in each subgroup stratified by quintiles.

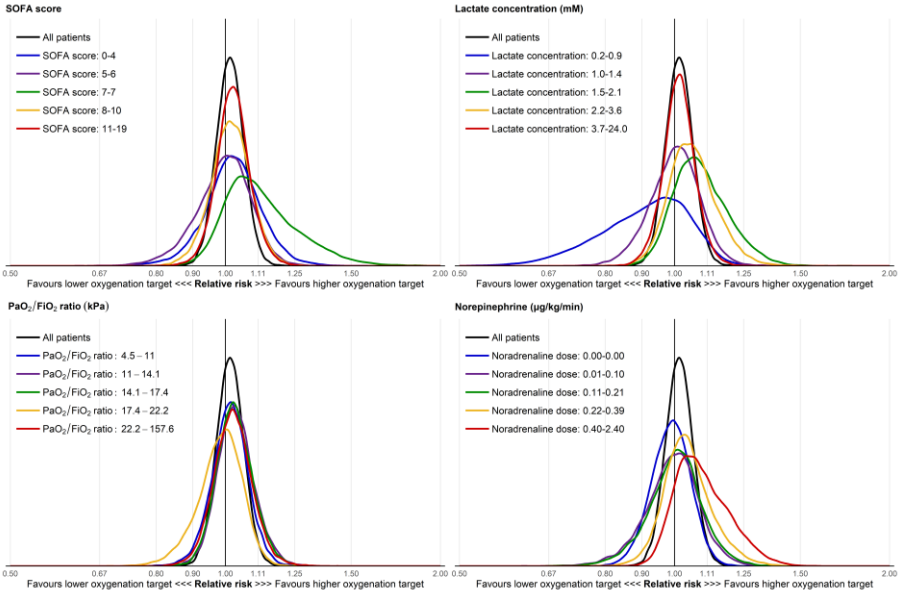


Figure 14. Subgroup-based plots for 90-day all-cause mortality. Posterior probability distributions of the relative risk (RR) of the treatment effect on 90-day all-cause mortality displayed according to Sequential Organ failure Assessment (SOFA) score, plasma lactate concentration, partial pressure of arterial oxygen to fraction of inspired oxygen (PaO_2/FiO_2) ratio, and norepinephrine dose at baseline (using weakly informative priors). Published with permission from the journal.[133]

4.2.4. Continuous HTE analyses

Conditional effects plots are presented in Figure 15, and illustrate the estimated interactions between oxygenation target allocation and 90-day all-cause mortality conditional on the baseline parameter in focus.

Our analyses found a 95% probability of a positive interaction between increasing baseline norepinephrine dose and the lower oxygenation target. This corresponds to a potential harmful effect (i.e. increased mortality) of a lower oxygenation target with increasing dose of continuously infused norepinephrine at baseline.

We found an 86% probability of a positive interaction with baseline plasma lactate concentration and the lower oxygenation target on mortality, corresponding to potential increased risk of death of the lower oxygenation target for patients with higher concentrations of plasma lactate at baseline.

We found a 65% probability for a positive interaction (i.e. potentially increased mortality risk) between the lower oxygenation target and increasing baseline SOFA scores (i.e. higher degree of organ failure).

We found a 76% probability for a positive interaction (i.e. potentially increased mortality risk) between the lower oxygenation target and decreasing baseline PaO_2/FiO_2-ratios (i.e. greater severity of respiratory failure).

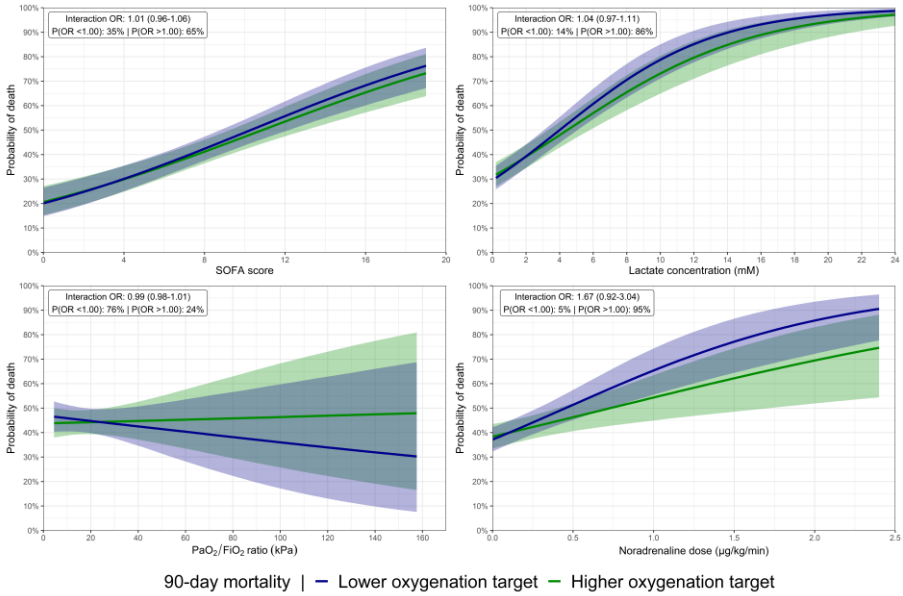


Figure 15. Conditional effects plots for 90-day all-cause mortality (weakly informative priors). Within each subplot is the OR (with 95% CrI) for the interaction-effect between the lower oxygenation target and the assessed baseline variable. SOFA score denotes sequential organ failure assessment score, PaO_2/FiO_2 ratio of partial pressure of arterial oxygen to fraction of inspired oxygen. Published with permission from the journal.[133]

4.2.5. Sensitivity analyses

Two sets of sensitivity analyses were conducted: one using *sceptic* priors; and one using *evidence-based* priors as defined in the protocol for the sudy.[96] Results from these analyses were in line with the primary analyses using *weakly informative* priors and are presented in the supplement for Paper II (Appendix B).[133]

4.3. Updated Cochrane review (Paper III)

4.3.1. Literature search

In the previous version of this review a total of 32,813 records were screened.[30] The same search string has also been updated and used for an additional review focussing on acutely ill patients (i.e. also including patients not admitted to the ICU).[159] For this updated version, an additional 10,954 titles and abstracts were screened for inclusion, resulting in a total of 46,323 records. Seven new trial reports were identified,[82–84,132,160–162] and one previously identified[71] was excluded due to overlap in patient population with a new report.[83] This left us with 16 reports for the qualitative synthesis[65,66,83,84,132,160–162,67–70,72–74,82] and 14 reports for the quantitative analyses,[65,66,160–163,67,69,70,74,82–84,132] as two trials did not report on any of our pre-defined outcomes.[72,73] An additional 10 relevant ongoing trials were identified. For additional details the reader is referred to Paper III (Appendix C).[134]

4.3.2. Characteristics of identified trials

Patients and setting

In total, 6,486 patients were randomised in the 16 identified trials. All trials restricted inclusion to adults (≥18 years). Three trials had specific age requirements: 40-70 years,[72] 18-65 years,[73] or 'elderly' (not specified).[160] The number of patients randomised ranged from 34[161] to 2,928,[132] the approximate mean age was 61 years, and the approximate mean proportion of men was 65%.

Patients were admitted to either a medical ICU, a surgical ICU, or a multidisciplinary (mixed) ICU before randomisation. Half of the identified trials (8/16) restricted inclusion to adults receiving IMV.[66,68–70,73,74,82,161]

Key characteristics of the included trials are summarised in Table 8.

Table 8. Summary of key characteristics of included trials

Author, year	Setting	Population	n	Respiratory failure	Results
Gomersall, 2002[67]	Medical ICU	Acute exacerbation of COPD	36	App. mean PaO_2 5.7 kPa	No differences in clinical outcomes
Mazdeh, 2015[72]	Medical ICU	Stroke	51	-	No difference in mortality or Barthel index, reduced disability in high oxygenation group at 6 months (mRS)
Taher, 2016[73]	ICU, not specified	Traumatic brain injury	68	-	Reduced disability in higher oxygenation group at six months (GOS, Barthel index, and mRS)
Panwar, 2016[68]	Mixed ICU	IMV	104	App. mean PaO_2/FiO_2 ratio 33.0 kPa	No differences in clinical outcomes. Feasible to implement conservative oxygenation strategy.
Girardis, 2016[65]	Mixed ICU	Expected ICU-stay ≥72 hours	480	54.6% respiratory failure, 67.1% MV	Increased hospital mortality in high oxygenation group (RR 1.33, 95% CI 1.00-1.78), increased risk of shock, liver failure, and bacteraemia in high oxygenation group.
Asfar, 2017[66]	Mixed ICU	IMV with refractory septic shock	442	App. mean PaO_2/FiO_2 ratio 29.9 kPa	No difference in mortality
Ishii, 2018[74]	Surgical ICU	IMV >12 hours	44	-	No difference in atelectasis formation
Lång, 2018[69]	Surgical ICU	Traumatic brain injury	65	-	No difference in mortality
Jakkula, 2018[70]	Medical ICU	Witnessed out of hospital cardiac arrest and IMV	123	-	No differences in clinical outcomes

Jun, 2019[160]	ICU, not specified	Acute exacerbation of COPD AND AMI not accepting PCI	87	-	Lower 14-day mortality rate, lower occurrence of malignant arrhythmia, lower rate of myocardial infarction, and higher extubation rate in high oxygenation group.
Yang, 2019[162]	Mixed ICU	Expected ICU-stay ≥72 hours	214	54.2% respiratory failure, 83.6% MV	No difference in mortality
Barrot, 2020[82]	Mixed ICU	IMV for ARDS <12 hours	205	Approximate mean PaO_2/FiO_2 ratio 15.8 kPa	Lower 90-day mortality in high oxygenation group (RR 0.68, 95% CI 0.47-0.99), increased incidence of intestinal ischaemia in low oxygenation group
Mackle, 2020[83]	Mixed ICU	Expected IMV or NIV beyond next calendar day	1,000	App. mean PaO_2/FiO_2 ratio 33.6 kPa	No differences in clinical outcomes
Martin, 2021[161]	ICU, not specified	IMV and NIV, and expected ≥72 hours	34	App. mean PaO_2/FiO_2 ratio 26.5 kPa	No differences in clinical outcomes
Schjørring, 2021[132]	Mixed ICU	Acute hypoxaemic respiratory failure	2,928	App. mean PaO_2/FiO_2 ratio 15.7 kPa, 58.6% MV	No differences in clinical outcomes
Gelissen, 2021[84]	Mixed ICU	≥2 positive SIRS-criteria	574	App. mean PaO_2/FiO_2 ratio 26.2 kPa	No differences in clinical outcomes

ARDS denotes acute respiratory distress syndrome, CI confidence interval COPD chronic obstructive pulmonary disease, FiO_2 fraction of inspired oxygen, GOS Glasgow outcome scale, ICU intensive care unit, kPa Kilo Pascal, IMV invasive mechanical ventilation, mRS modified Ranking Scale, NIV non-invasive mechanical ventilation, PaO_2 partial pressure of arterial oxygen, SIRS systemic inflammatory response syndrome, RR relative risk.

Oxygenation strategies in identified trials

Thirteen of the 16 trials included reported on all-cause mortality.[65,66,132,161,162,67–70,72,82–84] Eight trials were judged to be at overall low risk of bias for this outcome.[66,68,70,82,83,132,161,162] The identified trial used a wide range of different oxygenation strategies to define both intervention and control groups. These are summarised in Table 9.

Table 9. Interventions applied in RCTs on targeted oxygenation therapy in the ICU

Author, year	Higher group			Lower group		
	FiO_2	PaO_2	SaO_2/SpO_2	FiO_2	PaO_2	SaO_2/SpO_2
Gomersall, 2002[67]		>9.0 kPa			>6.6 kPa	
Mazdeh, 2015[72]	0.50			Supplemental oxygen not used		
Taher, 2016[73]	0.80			0.50		
Panwar, 2016[68]			≥96%			88-92%
Girardis, 2016[65]	≥0.40	≤20 kPa	97-100%		9.3-13.3 kPa	94-98%
Asfar, 2017[66]	1.00					88-95%
Ishii, 2018[74]	1.00				13.3 kPa	
Lång, 2018[69]	0.70			0.40		
Jakkula, 2018[70]		20-25 kPa			10-15 kPa	95-98%
Jun, 2019[160]	0.40-0.70			0.30-0.50		
Yang, 2019[162]	≥0.30		96%-100%			90-95%
Barrot, 2020[82]		12-14 kPa	≥96%		7.3-9.3 kPa	88-92%
Mackle, 2020[83]	FiO_2 <0.30 discouraged					90-96%
Martin, 2021[161]			≥96%			88-92%
Schjørring, 2021[132]		12 kPa			8 kPa	
Gelissen, 2021[84]		8-12 kPa			14-18 kPa	

FiO_2 denotes fraction of inspired oxygen, PaO_2 partial pressure of arterial oxygen, SaO_2 arterial oxygen saturation, SpO_2 peripheral oxygen saturation. Trials listed after publication year. Published with permission from the journal.[134]

4.3.3. Co-primary outcomes

All-cause mortality

Meta-analysis indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies both in trials at overall low risk of bias (MH random RR 0.99, 95% CI 0.91-1.09; I^2 = 13%; 4,945 participants; 8 trials; low certainty) (Figure 16) and in all included trials (MH random RR 1.01, 95% CI 0.94-1.10; I^2 = 9%; 5,973 participants; 13 trials; very low certainty) (Figure 17).

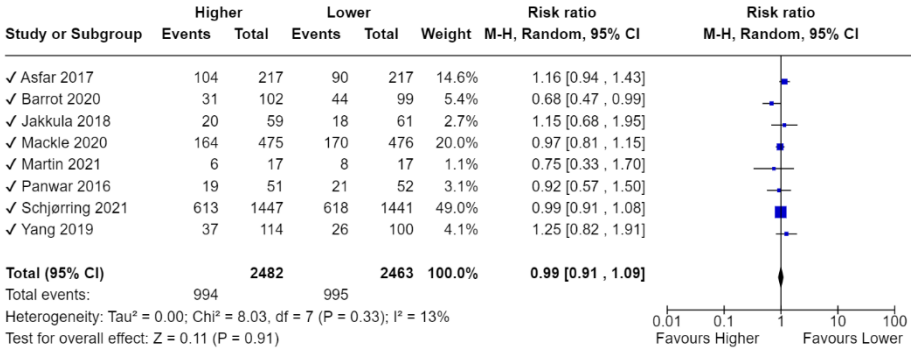


Figure 16. Meta-analysis of all-cause mortality at maximum follow-up in trials judged to be at overall low risk of bias. M-H, Random denotes Mantel-Haenszel random effects model, CI confidence interval. Published with permission from the journal.[134]

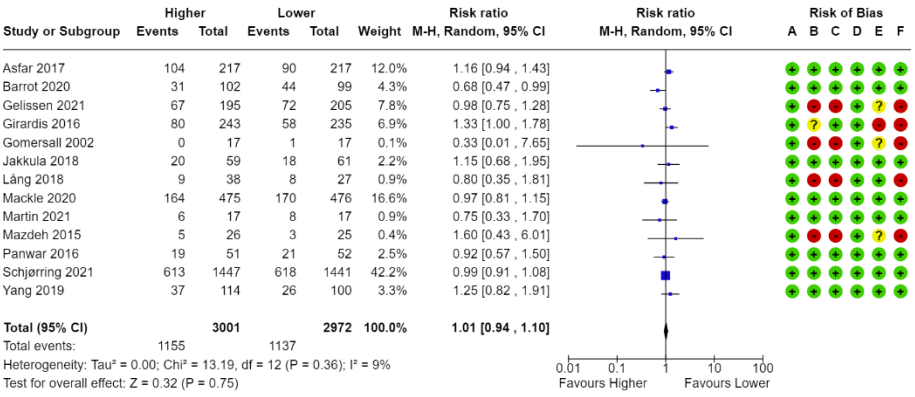


*Figure 17. Meta-analysis of all-cause mortality at maximum follow-up in all identified trials. M-H, Random denotes Mantel-Haenszel random effects model, CI confidence interval. For the domains of Risk of Bias: **A** denotes bias arising from the randomisation process; **B** risk of bias due to due to deviations from intended interventions; **C** risk of bias due to missing outcome data, **D** bias in measurement of the outcome, **E** bias in selection of the reported result, and **F** the overall bias. Green marker ● = 'low risk of bias', yellow marker ? = 'some concerns', and red marker ● = 'high risk of bias'. Published with permission from the journal.*[134]

Trial Sequential Analysis (TSA) could reject a relative risk increase (RRI) of 10% or more (Figure 18), equivalent to rejecting an absolute increase of 3.8 percentage points. The TSA CI for the intervention effect on the relative scale, adjusted for multiple outcomes, sparse data, and repetitive testing was 0.91 to 1.12.

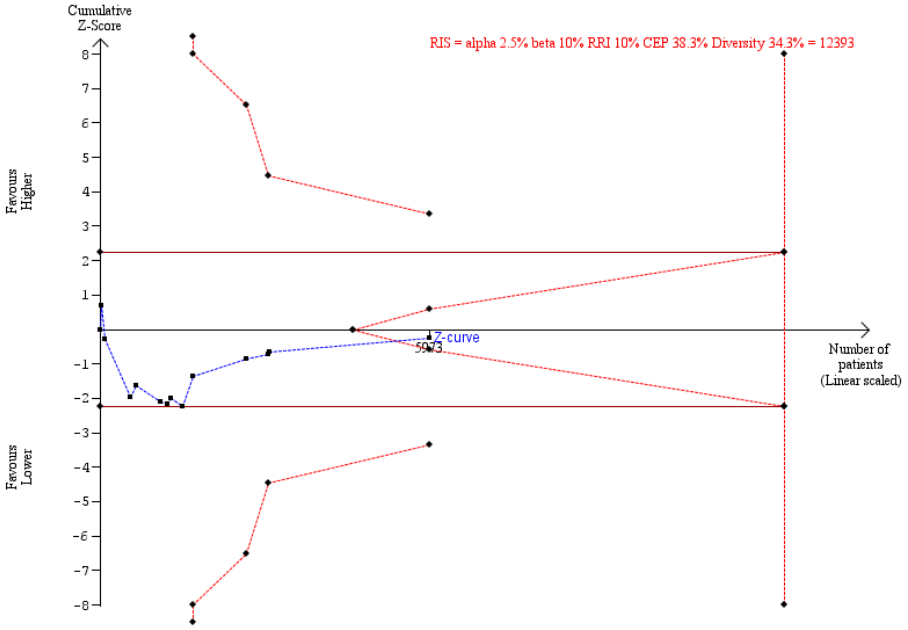


Figure 18. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the risk of all-cause mortality. The analysis was based on a mortality in the control group (control event proportion = CEP) of 38.3%, a relative risk increase (RRI) of 10%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 34.3%. Required information size (RIS) = 12,393. The cumulative Z-curve crossed the trial sequential monitoring boundary for futility. Published with permission from the journal.[134]

Proportion of patients with one or more serious adverse event

Three of the 16 included trials reported on the proportion of patients with one or more SAE as a composite outcome.[66,84,132] Two were judged to be at overall low risk of bias for this outcome.[66,132]

Meta-analysis indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies both in trials at overall low risk of bias (MH fixed RR 1.07, 95% CI 0.99-1.15; I^2 = 6%; 3,344 patients; 2 trials; low certainty) and in all

included trials (MH fixed RR 1.05, 95% CI 0.98-1.13; I^2 = 27%; 3,744 patients; 3 trials; very low certainty).

TSA could reject a relative risk increase (RRI) of 15% or more (Figure 19), equivalent to rejecting an absolute increase of 6.2 percentage points. The TSA CI for the intervention effect, adjusted for multiple outcomes, sparse data, and repetitive testing was 0.96 to 1.16.

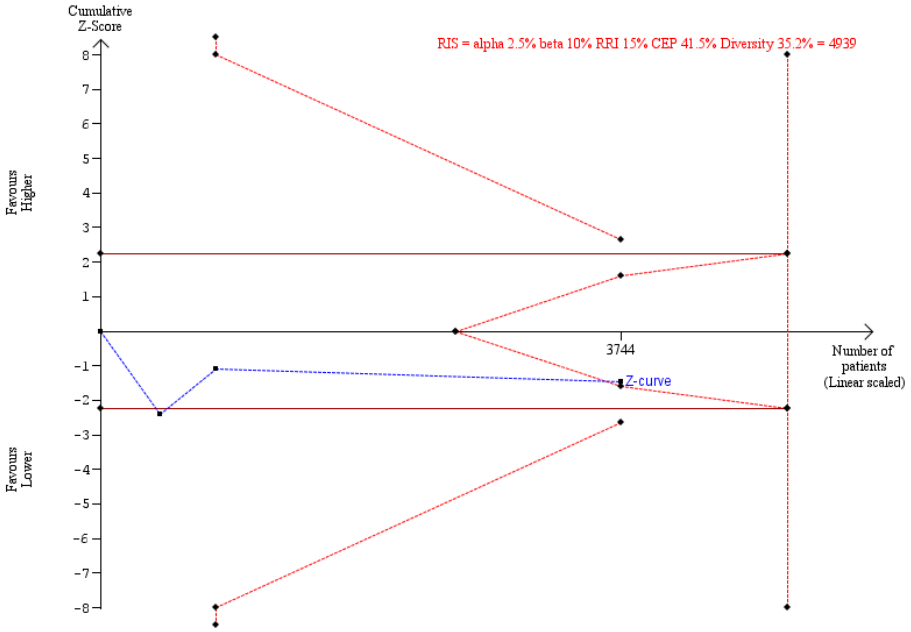


Figure 19.Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the proportion of patients with one or more serious adverse events. The analysis was based on a proportion in the control group (control event proportion = CEP) of 41.5%, a relative risk increase (RRI) of 15%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 35.2%. Required information size (RIS) = 4,939. The cumulative Z-curve crossed the trial sequential monitoring boundary for futility.

As sensitivity analyses, and since 14 of the 16 included trials[65,66,160–162,67–70,82–84,132] reported on the occurrence of any SAE according to the ICH-GCP definition,[164] two additional analyses on this matter were conducted: 1) the highest proportion of any specific SAE; and 2) the cumulated number of SAEs.

Eight of the 14 trials were judged to be at overall low risk of bias for this outcome.[66,68,70,82,83,132,161,162] In trials judged to be at overall low risk of bias, meta-

analysis of the highest proportion specific SAEs indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (MH fixed RR 1.00, 95% CI 0.94-1.06; I^2 = 42%; 4945 participants; 8 trials). Meta-analysis of all included trials found a similar result (MH fixed RR 1.00, 95% CI 0.95-1.06; I^2 = 38%; 6,031 patients).

In trials judged to be at overall low risk of bias, meta-analysis of the cumulated number of SAEs indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (MH fixed RR 1.02, 95% CI 0.98-1.06; I^2 = 0%; 4212 participants; 5 trials). Meta-analysis of all included trials demonstrated a similar result (MH fixed RR 1.03, 95% CI 1.00-1.06; I^2 = 67%; 6,053 patients; 14 trials).

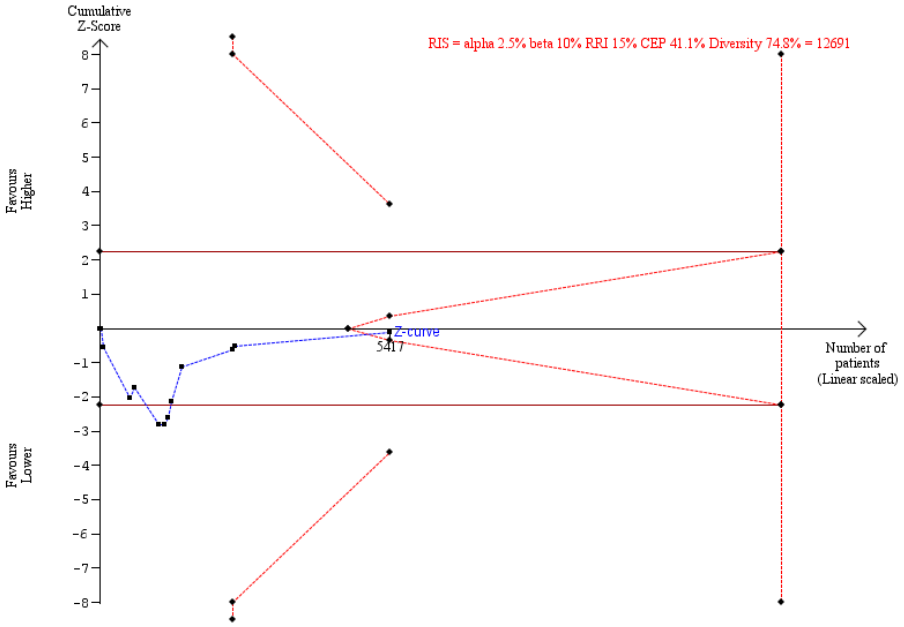


Figure 20. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the highest proportion of serious adverse events in all trials. The analysis was based on a proportion in the control group (control event proportion = CEP) of 41.3%, a relative risk increase (RRI) of 15%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 74.8%. Required information size (RIS) = 12,691. The cumulative Z-curve crossed the trial sequential monitoring boundary for futility.

TSA of the highest proportion of any specific SAE (Figure 20) and of the cumulated number of SAEs (Figure 21) in all trials rejected a relative risk increase (RRI) of 15% or more in both instances, as the trial sequential monitoring boundaries for futility were crossed. This was equivalent to rejecting absolute increases of 6.2 and 10.4 percentage points, respectively.

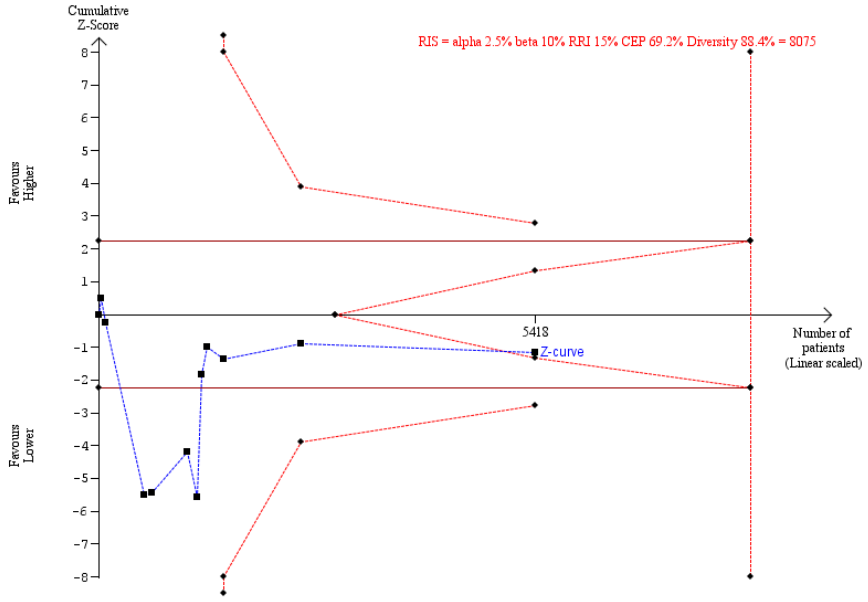


Figure 21. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the cumulated number of serious adverse events in all trials. The analysis was based on a proportion in the control group (control event proportion = CEP) of 68.0%, a relative risk increase (RRI) of 15%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 88.4%. Required information size (RIS) = 8,075. The cumulative Z-curve crossed the trial sequential monitoring boundary for futility.

Quality of life

Only one of the 16 included trials reported on 'quality of life' using any measure.[83] The trial recorded the mean reported health state scores (± SD) in survivors at 180 days after randomisation using the EuroQoL visual analogue scale (EQ-VAS) score.[135] However, this trial was judged to be at overall high of bias for this outcome as data was only available for 499 of 617 eligible patients. The means (±SD) in the higher and lower groups were 67.6 points (±22.4) and 70.1 points (±22.0), respectively; the mean difference was -2.5 points (95% CI -6.4-1.4; p = 0.22; 499 patients).[83] Certainty of evidence was very low.

4.3.4. Secondary outcomes

Lung injury

None of the identified trials reported any data on the occurrence of 'lung injury' defined as a composite outcome as specified in the review protocol (occurrence of ARDS, pulmonary fibrosis, or pneumonia).[149] Seven of the 16 included trials reported on the occurrence of specific lung outcomes: three trials reported on the occurrence of ARDS;[68,70,84] three trials reported on the occurrence pneumonia;[65,66,82] one trial reported on both ARDS and pneumonia;[69] but no trial reported on the occurrence of pulmonary fibrosis. Three of the seven trials were judged to be at overall low risk of bias for this outcome.[68,70,82]

As with SAEs, both the 'highest proportion' and 'cumulated number' of lung injuries (as a composite outcome) were evaluated.

Meta-analysis in trials judged to be at overall low risk of bias of the highest proportion of patients with lung injury indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (MH random RR 1.16, 95% CI 0.74-1.81; I^2 = 0%; 424 participants; 3 trials). Meta-analysis of all trials found a similar result (MH fixed RR 1.06, 95% CI 0.82 to 1.36; I^2 = 0%; 1,942 patients; 7 trials; very low certainty).

Meta-analysis in trials judged to be at overall low risk of bias of the cumulated number of patients with lung injury indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (MH random RR 1.16, 95% CI 0.74-1.81; I^2 = 0%; 424 participants; 3 trials). Meta-analysis of all trials found a similar result (MH random RR 1.02, 95% CI 0.80-1.31; I^2 = 0%; 1,942 patients; 7 trials; very low certainty).

Myocardial infarction

Three of the 16 included trials reported on the occurrence of myocardial infarction,[84,132,160] but only one was judged to be at overall low risk of bias for this outcome.[132]

Meta-analysis of all trials indicated no evidence of a difference between higher or lower oxygenation strategies (MH random RR 0.59, 95% CI 0.25-1.38; I^2 = 17%; 3,368 patients; 3 trials; very low certainty).

Stroke

Four of the 16 included trials included reported on the occurrence of stroke,[82–84,132] and two trials were judged to be at overall low risk of bias for this outcome.[82,132]

Meta-analysis of trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (MH fixed RR 1.04, 95% CI 0.59 to 1.83; I^2 = 49%; 3,111 patients; 2 trials; very low certainty).

Meta-analysis of all included trials found a similar result (MH fixed RR 1.12, 95% CI 0.65-1.92; I^2 = 0; 4,476 patients; 4 trials; very low certainty).

Sepsis

Two of the 16 trials included reported on the occurrence of sepsis,[65,82] but only one trial was judged to be at overall low risk of bias for this outcome.[82]

Meta-analysis of all included trials indicated evidence of benefit from lower oxygenation strategies as compared with higher (MH random RR 1.81, 95% CI 1.11-2.95; I^2 = 0%; 646 patients; 2 trials; very low certainty).

5. DISCUSSION

At day 90 after randomisation in the HOT-ICU trial, 618 of 1,441 patients in the lower group (42.9%) and 613 of 1,447 in the higher group (42.4%) had died: adjusted RR 1.02 (95% CI 0.94-1.11).[132] In the secondary, Bayesian analysis of the trial we estimated the adjusted RR for 90-day all-cause mortality to most likely to be 1.02 and with 95% probability between 0.93 and 1.11, thus in line with the primary, frequentist approach when considering the overall ('marginal') effect.[133] This was not surprising, since the amount of data in the trial, as expected, overwhelmed the statistical model, and thus dominated the posterior probability distribution. The two analyses yielded similar point estimates for the intervention effect on 90-day all-cause mortality, suggesting only a very small probability of harm from the lower oxygenation target as compared to the higher. Thus, the HOT-ICU trial did not demonstrate any difference in the mortality effect, when comparing a lower and a higher oxygenation target in adults acutely admitted to the intensive care unit with acute hypoxaemic respiratory failure. Neither did we find any differences in any of the secondary outcomes, and with comparable risks of serious adverse events (adjusted RR 0.95, 95% CI 0.84-1.07), it appeared equally safe to target PaO_2 of 8 or 12 kPa.

As pointed out in the introduction of this thesis, '*absence of evidence is not evidence of absence*',[125] and the results from the HOT-ICU trial, though overall neutral, did not preclude the existence of clinically important differences in effects of the two oxygenation targets applied. Based on the 95% CI for all-cause mortality, a lower oxygenation target could result in both a 6% relative risk reduction and an 11% increase, as compared with a higher target, though effect sizes closer to the point estimate are more compatible with the data as previously mentioned (see section 1.5). Also, the CrI from the Bayesian analysis included both harmful and beneficial effects of a lower oxygenation target, but relative effect sizes on mortality larger than 20% (as a priori specified in the HOT-ICU protocol)[123] were highly unlikely, though smaller effects were still probable. Given the high mortality rate in the investigated population (42.6% overall at 90 days), even small effects are relevant, and thus more information is required to ascertain this issue. In the updated Cochrane review, we identified 16 RCTs, randomising a total of 6,486 patients, reporting on the effects of higher versus lower oxygenation strategies in critically ill patients admitted to the ICU. Notably, by using Trial Sequential Analysis, we could reject a 10% relative risk reduction or increase for all-cause mortality, and 15% for the proportion of patients with one or more SAE (including the 'highest reported proportion' and 'estimated cumulated number') at maximum follow-up. The 15% cut-off when considering SAEs is not reflected in the conclusion of the review, as a level of 20% was pre-defined, and we included 10% as a sensitivity analysis, as described in the section 'Differences between protocol and review' in Paper III

(Appendix C).[134] The conclusion of the review was that the true effects of higher versus lower oxygenation strategies in adult patients admitted to the ICU were still uncertain, due to low or very low certainty of evidence. The overall findings of this review are in line with other recent systematic reviews on this matter,[32,33,159] whilst other reviews report that higher oxygenation strategies may be harmful or not beneficial.[79,165,166] When considering acutely ill patients in general (i.e. both patients admitted to an ICU and those not), the most recent review on this matter concluded that a 15% relative change in mortality and 20% in occurrence of SAEs could be rejected.[159] However, the collected evidence was inconclusive with regards to smaller effects due to limited data, and the authors call for additional trials on this matter.

At the moment no minimal clinically relevant difference for mortality has been established, as one may argue that any difference is important. When assessing minimally clinical relevant differences it is generally recommended to use effects sizes defined on the absolute scale, as effects on the relative scale depend on the baseline risk.[167,168] Probabilities for a range of effects sizes for 90-day all-cause mortality on the absolute scale are presented in Paper II (Appendix B).[133] At the moment, two major RCTs on targeted oxygenation therapy in the ICU, the MEGA-ROX and UK-ROX trials, are being conducted.[169,170] The MEGA-ROX trial is an extension of the ICU-ROX trial,[83] based in Australia and New Zealand, and plans to recruit 40,000 mechanically ventilated ICU patients. The UK-ROX trial is based in the United Kingdom and plans to recruit 16,500 mechanically ventilated ICU patients from 100 ICUs. The MEGA-ROX and UK-ROX trials are designed to assess absolute mortality reductions in critically ill patients in the ICU with a lower oxygenation strategy, as compared to a higher strategy, of 1.5 and 2.5 percentage points, respectively. For the sake of argument, the minimal clinically important mortality difference could be defined arbitrarily as low as an absolute risk increase or reduction of 1 percentage point, equivalent to 1,000 lives saved for every 100,000 patients treated. The estimation of probabilities for many different effect sizes is one of the major advantages with the Bayesian approach, as the posterior probability distribution *is* the result of the analysis. Thus, no issues with multiple comparison are relevant, as estimation of probabilities of different effect sizes is a matter of integrating the area under the curve of the posterior distribution. As such, we found that the probability of an RD ≥1 percentage point (i.e. 'benefit of the higher oxygenation target') was 42.1%, and was 18.8% for an RD ≤1 percentage point (i.e. 'benefit of the lower oxygenation target', in the HOT-ICU trial. Thus, the probability of the effect being between these two margins (i.e. 'no clinically important difference') was 39.1%. Based on this, we remain uncertain about the effects of targeted oxygenation in the HOT-ICU trial, though data still point to an overall neutral result.

Perhaps the principal reason that we did not find any difference between a higher or a lower oxygenation strategy in the identified RCTs in the updated review could be

that patients in most trials, neither achieved sufficient hypoxaemia nor hyperoxaemia to substantially impact overall mortality risk. Alternatively, targeted oxygenation in critically ill patients could be regarded as 'equally safe', as long as one aims within the 'normal range' of oxygenation. This range has, in healthy individuals, been suggested to be 10.7-13.3 kPa at sea level.[171] The levels of hypoxaemia and hyperoxaemia are not unequivocally defined, but the most widely used definitions are a PaO_2 below 8 kPa or an SaO_2 less than 90%, or a PaO_2 above 16 kPa, respectively.[10] The concept of 'safe margins of oxygenation' has been proposed by several,[9,10,36] and most elegantly illustrated by Pastene and Leone, as depicted in Figure 22.[172] Here, a range of oxygenation targets may safely apply to a number of different clinical circumstances with equal effects on outcomes, and are generally considered 'safe' for most patients. On the one end of the oxygenation spectrum anaerobe metabolism and cellular death ensues due to hypoxia, whilst on the other end excessive ROS productions occurs due to hyperoxia – both causing harm. However, the transition zones (marked with '?' in Figure 22) are currently not well established in either direction.

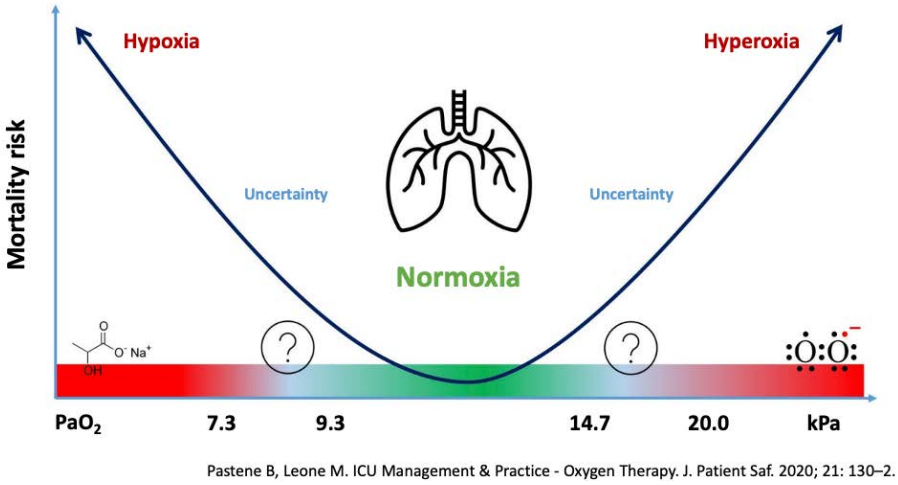


Figure 22. The uncertain effect of the arterial partial pressure of oxygen (PaO_2) on mortality risk. Adapted from Pastene and Leone and published with permission from the journal.[172]

In this relation, it is important to note that only the HYPERS2S trial investigated the effects of evident hyperoxaemia, as patients in the intervention group received an FiO_2 of 1.00 while invasively mechanically ventilated for the first 24 hours.[66] The trial suggested benefit in patients with septic shock of a lower oxygenation target with lowered mortality at both 28 and 90 days after randomisation, though these findings

were not statistically significant. In addition, hyperoxaemia appeared to increase the occurrence of atelectases, and ICU acquired weakness.

The results from the OXYGEN-ICU trial suggested that a liberal oxygenation strategy allowing a PaO_2 up to 20 kPa (i.e. within the proposed hyperoxaemic range) was harmful to patients, both in terms of higher mortality (within the ICU and in-hospital), but also in terms of higher incidences of shock, liver failure and bacteraemia.[65] Albeit not statistically significant, the $LOCO_2$ trial also suggested decreased rates of new 'septicaemia' in the lower oxygenation group.[82] Due to practical concerns, we did not perform any registration of infections, positive blood cultures, or similar measures in the HOT-ICU trial. We were therefore not able to quantify the occurrence of new infectious episodes or sepsis per se, although this would have been clinically relevant. We did however find similar proportions of patients with new episodes of shock in both groups. However, as 'new shock' was not defined as an individual outcome the difference was not tested statistically. Meta-analysis in the updated Cochrane review suggested a benefit of a lower oxygenation strategy on the occurrence of new episodes of sepsis (fewer episodes), as compared to a higher oxygenation strategy, though with very low certainty evidence.[134]

In August 2021, Gelissen et al. published the results from their 'Optimal Oxygenation in the Intensive Care Unit' (O_2-ICU) trial.[84] This was a Dutch, multicentre trial conducted from February 2015 to January 2019. Adult ICU patients with systemic inflammation and with an expected ICU stay of at least 48 hours, were randomised to either a 'low-normal oxygenation strategy' (PaO_2 8-12 kPa) or a 'high-normal oxygenation strategy' (PaO_2 14-18 kPa) until a maximum of 14 days in the ICU. A total of 574 patients were randomised, but only 400 were included in the intention-to-treat cohort: 173 due to lack of consent; and one patient was excluded post-randomisation due to development of severe ARDS in the high-normal group. Included patients were primarily admitted to the ICU due to acute medical conditions, with systemic infections or pneumonia being the dominant diagnoses. At baseline, the median PaO_2/FiO_2 ratio was approximately 26.4 kPa, corresponding to mild-to-moderate ARDS, and similar to that found in the HYPERS2S trial. The investigators achieved their intended target in the low-normal group (mean PaO_2 10.8 kPa, IQR 9.8-12.0 kPa), but not in the high-normal group (median PaO_2 12.8 kPa, IQR 10.9-14.9 kPa). Main outcome was the novel, but difficult to interpret, non-respiratory cumulative daily delta Sequential Organ Failure Assessment score ('$SOFA_{rank}$'), which has not been used in any other comparable RCT to date. No significant differences were demonstrated for any outcomes, including mortality. Despite a statistically significant separation in PaO_2 between the two groups, there were still substantial overlap in interquartile ranges during the entire intervention period, perhaps contributing to the neutral trial result. In contrast, no overlap of IQRs for PaO_2 was found in the HOT-ICU trial.

Most recently, in September 2021, a small-scale Chinese RCT investigating 'conservative' versus 'conventional' oxygen therapy in mechanically ventilated patients with pneumonia, was published.[173] The trial is not included in the Cochrane review as it was published after the literature search was run, and the trial does not appear in any available clinical trial registries. We were therefore unable to identify this trial in a timely manner. Only an abstract could be retrieved, despite the corresponding author has been contacted with request for a copy of the published report (which is in Chinese). In all, 51 patients were randomised to 'conventional' oxygen therapy (PaO_2 > 150 mmHg (20 kPa) or SpO_2 >96%) and 55 were allocated to 'conservative' oxygen therapy (PaO_2 75-100 mmHg (9.3-13.3 kPa) or SpO_2 90-92%). The investigators report significantly lower ICU mortality, longer duration of mechanical ventilation, lower incidences of nosocomial bloodstream infection, liver insufficiency, shock, and decreased ICU acquired weakness in the conservative oxygen therapy group as compared to the conventional. Inclusion of the reported data in an updated meta-analysis on all-cause mortality at longest follow-up (with a higher target being the intervention) did not alter the result markedly (updated random effects RR: 1.03, 95% CI 0.94-1.13; I^2 = 21.4%, 6079 participants; 14 RCTs) (unpublished data). Despite reporting a significant difference in oxygenation, neither group achieved their designated oxygenation target; the PaO_2 was reported as 68.9±4.7 mmHg (9.2±0.6 kPa) in the conservative group and 75.2±6.0 mmHg (10.0±0.8 kPa) in the conventional group. The trial is severely underpowered to demonstrate the reported difference in ICU mortality (27.3% in the conservative versus 45.1% in the conventional group), as this would require 228 patients with a power of 80%, and 304 if a power of 90% was instituted (both at an alpha of 5%). As no full-text has been available, it is very difficult to evaluate risk of bias. Also, no baseline characteristics are reported, and further comparison with the other identified RCTs is thus challenging. The findings reported from this trial must consequently be interpreted with caution.

The remaining identified RCTs, including the HOT-ICU trial, have all tested the effects of two oxygenation strategies within the relative normoxic range.

Mackle et al. reported in October 2019 (online ahead of print) the results from their 'Intensive Care Unit Randomised Trial Comparing Two Approaches to Oxygen Therapy' (ICU-ROX) trial.[83] This was a multicentre trial conducted in Australia and New Zealand from September, 2015 to May, 2018. At completion, the investigators had recruited the planned 1,000 mechanically ventilated adult ICU patients who were expected to receive mechanical ventilation (IMV and NIV) beyond the next calendar day. Patients were randomised to either 'conservative oxygen therapy' (SpO_2 90-96) or 'usual care' (SpO_2 ≥91%, FiO_2 ≥0.30) for the first 28 days of ICU admission. Around one third of patients were admitted after surgery (primarily after emergency surgery), 40% had acute brain disease, roughly 17% had suspected hypoxic-ischaemic encephalopathy, whereas only 14% of patients were admitted with a respiratory illness. This was also reflected in the baseline mean PaO_2/FIO_2 ratio

of 33.6 kPa, equivalent to mild ARDS, suggesting that a large proportion of patients most likely did not have hypoxaemic respiratory failure at ICU admission. In the main cohort the trial was neutral in its outcomes, but the investigators found a potential benefit of a conservative oxygenation strategy, in terms of statistically significant increased 28-day ventilator freedom (median difference 21.1 days, 95% CI 10.4-28.0 days), and lower 180-day all-cause mortality (RR 0.73, 95% CI 0.54-0.99), in patients admitted to the ICU with suspected hypoxic ischaemic encelopathy. In these patients, a (non-significant) signal for improved neurological outcome in the conservative oxygenation group at 180 days was also found.

In the March 2020 issue of the New England Journal of Medicine, which also carried the print version of the ICU-ROX trial, Barrot et al. reported the results from their 'Liberal Oxygenation Versus Conservative Oxygenation in ARDS' ($LOCO_2$) trial.[82] This was a French multicentre trial, conducted from June 2016 to September 2018. Adult ICU patients invasively mechanically ventilated due to ARDS were randomised to either a 'liberal oxygenation strategy' (PaO_2 12-14 kPa) or a 'conservative oxygenation strategy' (55-70 mmHg = 7.3-9.3 kPa) for the first seven days of ICU admission. The conservative strategy was based on the recommended oxygenation level in trials conducted by the ARDS-network (55-80 mmHg = 7.3-10.7 kPa),[37,56,174] and the $LOCO_2$ trial was essentially a 'proof of concept' trial for this strategy. Patients primarily presented with ARDS due to pulmonary causes, and had a mean PaO_2/FiO_2 ratio at baseline of 15.8 kPa. The investigators achieved a clear separation in oxygenation measures between the two groups. In the HOT-ICU trial, 58.6% of patients were invasively mechanically ventilated at baseline, and the overall median PaO_2/FiO_2 ratio was 15.7 kPa (about half of that found in the ICU-ROX trial), but similar to that of the $LOCO_2$ trial, corresponding to moderate-to-severe impairment of gas exchange at baseline according to the ARDS-categories.[82,83,132] The $LOCO_2$ trial was prematurely halted due to slow recruitment (after inclusion of 205 of 850 patients planned) and safety concerns arising from a non-pre-planned interim analysis which indicated increasing risk of intestinal ischaemia (non-pre-defined outcome) in the conservative group (five cases versus none in the conventional group). The overall incidence of intestinal ischaemia in the $LOCO_2$ trial was 2.5% (five cases among 201 randomised patients), and for comparison, the incidence of intestinal ischaemia in the HOT-ICU trial was 2.1% (61 cases among 2,910 randomised patients) with similar incidences in both groups.[82,132] At day 28 there was no significant difference in all-cause mortality in the $LOCO_2$ trial. However, the investigators found a significantly increased mortality risk in the conservative group at 90 days: mean increase 14.0 percentage points (95% CI 0.7-27.2 percentage points). Overall, the $LOCO_2$ trial suggested a beneficial effect from a more liberal oxygenation strategy among invasively mechanically ventilated ARDS patients as compared to a more conservative strategy.

Patients included in the ICU-ROX trial were slightly younger (approximate mean age 58 year)[83] than those included in the $LOCO_2$ (approximate mean age 63 year)[82] and

OXYGEN-ICU trials (approximate median age 64 years),[65] and substantially younger than those included in the O_2-ICU and HYPERS2S (median ages 68 years)[66,84] and HOT-ICU trials (median age 70 years).[132] Also, overall approximate mortality rates were lower in four of the trials compared to the HOT-ICU trial (42.6% at 90 days): OXYGEN-ICU 21.0% in-hospital; ICU-ROX 33.4% at 90 days; O_2-ICU trial 34.8% at 90 days; and $LOCO_2$ trial 37.3% at 90 days, but more similar to the findings of the HYPERS2S trial with 44.7% at 90 days. Since the HYPERS2S trial investigated patients with septic shock this is not surprising as this subgroup of ICU patients have a notoriously high mortality risk.[175] These facts, in combination with the vast number of clinical differences between the trials as discussed in Paper III (Appendix C),[134] could potentially limit comparison of findings across trials.

This leaves us with two RCTs suggesting a benefit from a lower oxygenation strategy,[65,66] two trials suggesting a benefit from a higher oxygenation strategy,[82,173] and three trials not demonstrating any significant differences.[83,84,132] However, the OXYGEN-ICU, HYPERS2S, and $LOCO_2$ trials were all stopped prematurely, were severely underpowered, as were the recent trial by Yang and Zhang, and thus all have high risks of reporting spurious findings.[65,66,82,173] In addition, the O_2-ICU trial had substantial loss to follow-up, thus diminishing its discriminatory power.[84]

Several other RCTs were identified in the updated Cochrane review, but have not been discussed presently due to their relatively small sizes or pilot-design.[67–73,161,162] For further details on all trials included in the updated Cochrane review see Paper III (Appendix C).[134]

5.1. Heterogeneous treatment effects

As was suggested by Hochberg et al., the balancing point between the risks of hypoxaemia and hyperoxaemia could be influenced by a range of clinical factors, including degree of illness and comorbidities, i.e. heterogeneous treatment effects depending on baseline characteristics.[176] This lends weight to individualisation of oxygen supplementation, and necessitates further investigations. Though the statistical diversity was low in most meta-analyses conducted in the updated Cochrane review, as e.g. measured by the I^2 or Chi^2 statistics, the underlying clinical diversity (i.e. clinically relevant differences between trials) could still be substantial, thus biasing the overall result.[177] However, it was beyond the scope of the review to statistically adjust for any clinical diversity in the outlined meta-analyses as the content was dictated by the published review-protocol.[149] We did, however, perform a range of pre-defined subgroup and sensitivity analyses to characterise part of this issue. When investigating heterogeneity via subgroup analyses one investigates the influence of a clinical characteristic on the intervention effect, which is therefore an assessment of statistical heterogeneity induced by clinical diversity. However,

clinical diversity may still exist without any significant statistical heterogeneity, and vice versa. We did not find any statistically significant effects on the investigated outcomes in any of the investigated subgroups.

Despite a neutral marginal mortality effect in the HOT-ICU trial, heterogeneous treatment effects (HTE) (i.e. 'conditional' effects), being non-random differences in the effect of oxygenation strategy on mortality *conditional* on baseline characteristics, could still be present.[178,179] We addressed this issue by using both frequentist sub-group analyses, and Bayesian HTE analyses.[132,133] In the primary publication of the HOT-ICU trial, the results from the following pre-defined subgroups were published: 1) patients with shock at randomisation; 2) patients receiving invasive mechanical ventilation at randomisation; 3) type of ICU admission (medical; elective surgical; emergency surgical); 4) patients with known COPD at randomisation; 5) patients with acute traumatic brain injury at randomisation; and 6) patients resuscitated from cardiac arrest. Overall, we did not find any statistically significant interaction between the selected baseline characteristics and oxygenation target allocation on 90-day all-cause mortality (see section 4.1.6 for results from the frequentist sub-group analyses).[132]

The Bayesian analyses of HTE in the HOT-ICU trial suggested increasing harm from the lower oxygenation target with increasing norepinephrine dose at baseline, implying that patients with increasing levels of circulatory insufficiency were less likely to die if their oxygenation levels were targeted at a PaO_2 of 12 kPa rather than 8 kPa.[133] As with the marginal effect estimate, none of the major RCTs on oxygenation strategies have explored HTE in a comparable manner.[65,66,82–84] However, our finding is to some extent supported by the results of the traditional frequentist sub-group analysis in the HOT-ICU trial of patients with shock at baseline (RR 1.07, 95% CI 0.95-1.21), as reported in section 4.1.6. The point estimate favours the higher oxygenation target, and most of the 95% CI also does. However, the two subgroups are not absolutely comparable as shock was defined as 'plasma lactate concentration at baseline ≥2 mM AND the use of vasopressors or inotropes', whilst in the Bayesian HTE analysis we considered the dose of norepinephrine at baseline and lactate concentration separately. No convincing signal for HTE for lactate concentration and oxygenation target allocation was suggested. Also, we did not find any convincing suggestions for HTE in any of the other subgrouping schemes nor on the continuous scale for any of the other selected baseline characteristics (PaO_2/FiO_2 ratio, and SOFA score). The ICU-ROX trial reported potential harm of a lower oxygenation strategy in a subgroup of patients with sepsis at baseline, even though the finding was not statistically significant.[180] A similar suggestions of increased mortality was found in the subgroup of patients with shock at baseline in the OXYGEN-ICU trials, but again this finding was not statistically significant.[65] Conversely, in the HYPERS2S trial, which restricted recruitment to patients with septic shock, hyperoxia as compared to normoxia increased all-cause mortality at both 28 and 90 days, albeit not statistically significant, but did significantly increase

the incidence of SAEs.[66] In a post-hoc subgroup analyses of the HYPERS2S trial, hyperoxia as compared to normoxia, increased mortality in patents with baseline lactate levels >2 mmol/L, but not in patients with lower lactate levels.[181] As only limited data on patients with sepsis exists, no recommendations on oxygenation in adults with sepsis-induced hypoxemic respiratory failure have yet been proposed by the Surviving Sepsis Campaign.[182] More studies on this subject are needed and are highly clinically relevant.[183]

In a post-hoc subgroup analysis of 166 patients with suspected hypoxic ischaemic encephalopathy at baseline in the ICU-ROX trial, the investigators reported interaction with oxygenation target allocation, favouring the lower strategy in terms of lower mortality at day 180 after randomisation (unadjusted RR 0.73, 95% CI 0.54-0.99; adjusted OR 0.56, 95% CI 0.25-1.23).[83,184] Hyperoxaemia in patients resuscitated from cardiac arrest was in a systematic review of observational studies also associated with increased mortality.[185] Young et al. also suggested potential benefit of a conservative oxygenation strategy among adult patients admitted to the ICU after cardiac arrest, in an individual-level patient data meta-analysis.[186] However, certainty of evidence was low or very low, especially due to differences of included studies and indirectness.[186] In contrast, the COMACARE trial, a 2^3 factorial RCT which randomised 120 patients resuscitated from out-of-hospital cardiac arrest to either normoxia (PaO_2 10-15 kPa) or moderate hyperoxia (PaO_2 20-25 kPa) (in addition to low-normal versus high-normal $PaCO_2$ and mean arterial pressure) found no difference in 30-day mortality.[70] However, this was a feasibility pilot trial and was not powered to assess any differences in hard end-points. The investigators did find that a higher oxygenation strategy resulted in higher levels of cerebral oxygen saturation, as compared to the lower oxygenation strategy, though this did not translate into any differences in any of the other outcomes reported in this trial. To enable comparison with the subgroup of patients in the ICU-ROX trial with suspected hypoxic-ischemic encephalopathy at baseline we also decided to include data on patients with cardiac arrest prior to randomisation. In these 332 patients we found no evident between-group difference in 90-day all-cause mortality though the point-estimate favours the higher oxygenation group (adjusted RR 1.09, 95% CI 0.92-1.28).[132] These findings underline the caution one must employ when considering results from subgroup analyses and observational studies.

In the subgroup of 562 patients diagnosed with COPD at baseline in the HOT-ICU trial, we found no significant difference in 90-day all-cause mortality (RR 0.98, 95% CI 0.82-1.17). This is the largest report on targeted oxygenation therapy in COPD patients in the ICU to date. In 2002, Gomersall et al. reported the results from their single-centre pilot RCT involving 36 patients with acute exacerbation of COPD.[67] Patients in this trial were randomised to target either a $PaO_2 > 9.0$ kPa or a $PaO_2 > 6.6$ kPa. Those in the lower group would also receive doxapram (a respiratory stimulant acting on the chemoreceptors in the carotid bodies) if acidosis (defined as $pH < 7.20$) developed, and in the higher group doxapram would be supplied in case of

symptomatic acidosis. No differences in outcomes were found between the two groups, and the findings of this trial have not had a great impact on clinical practice. At the moment, the generally recommended oxygenation level for acutely ill COPD patients is an SpO_2 of 88-92%,[9–11] despite the evidence for this recommendation being based in its essence on a single RCT by Austin et al. from 2010.[187] In this cluster-randomised trial, the investigators randomised 405 patients, with presumed acute exacerbation of COPD in the pre-hospital setting, to either titrated oxygen therapy (SpO_2 88-92%) or conventional high-flow oxygen (8-10 l/min via non-rebreather facemask) provided by paramedics. The intention-to-treat analyses of all patients and of those with post-randomisation confirmed COPD (n = 214) both demonstrated statistically significant mortality reduction if patients were allocated to titrated oxygen therapy as compared to the conventional approach: RR 0.42 (95% CI 0.20-0.89, p = 0.02) and RR 0.22 (95% CI 0.05-0.91, p = 0.04), respectively. The level of intended oxygenation in this trial corresponds to our lower target of a PaO_2 of 8 kPa. However, the COPD patients in the trial by Austin et al. reported a mean PaO_2 of 10.6±3.3 kPa and 13.1±6.1 kPa upon arrival to hospital in the titrated and conventional groups, respectively.[187] This was slightly more than the levels seen in the HOT-ICU trial overall. Additionally, the average time of titrated oxygen in the prehospital setting was a mere 45 minutes, in comparison to the maximum of 90 days in the HOT-ICU trial. These findings are supported by a recent retrospective cohort study, demonstrating the lowest mortality both among all included patients, but also for patients with normocapnia and hypercapnia, if SpO_2 was between 88-92%.[188] However, one-year mortality among patients surviving to hospital discharge was highest in this group if patients were normocapnic upon index admission (lowest mortality at SpO_2 97-100%), but comparable to other SpO_2 ranges if patients were hypercapnic.[188] Further investigation into the optimum oxygenation strategy for patients with COPD and acute hypoxaemic respiratory failure is highly relevant and warranted to inform future clinical practice.[189]

Among patients with traumatic brain injury, data in the HOT-ICU trial suggested a (albeit not statistically significant) benefit of the higher target (RR 2.63, 95% CI 0.55-12.54). However, this finding is most likely due to chance as we only recruited 22 patients with traumatic brain injury. The major reason for the low number of patients in this category is that we generally had great difficulty in engaging the neuro-intensive care units due to widespread concerns with the lower target. However, a recent meta-analysis of acutely ill patients (including those not admitted to an ICU) with any cerebral disease did not find any evidence of benefit from either a lower or higher oxygenation strategy.[159] This question therefore needs further investigation to produce firm evidence for this subgroup.

A total of 366 patients were in the HOT-ICU trial registered as diagnosed with ARDS at baseline by, judged by the clinicians according to the 2012 ARDS Berlin-criteria.[41] We did however note that 65 of these patients were on open oxygenation systems at baseline, and seven patients had PaO_2/FiO_2 ratios >40 kPa, which is not in line with

the criteria. Nevertheless, most included patients in the entire trial cohort had PaO_2/FiO_2 ratios below 40 kPa at baseline and remained below this level throughout their stay in intensive care (see section 4.1.3). The point estimate for the ARDS subgroup pointed to a potential benefit of the lower target (RR 0.88, 95% CI 0.71-1.10), which is in contrast to the findings from the $LOCO_2$ trial that suggested benefit from a higher oxygenation strategy among ARDS patients,[82] however, being of very low certainty as per a recent Cochrane review.[190] Girardis et al. also reported a potential benefit of a conservative oxygenation strategy among patients with respiratory failure at admission, in terms of lower ICU mortality (absolute risk reduction 12.8 percentage point, 95% CI 2.3-23.0, p = 0.02), but failed to report the criteria for this categorisation.[65] Modification of the 2012 ARDS consensus definition[41] for inclusion and exclusion criteria in RCTs on ARDS has been demonstrated to being common, may explain variations in reported control-mortality rates, with severity of hypoxaemic respiratory failure (measured by the PaO_2/FiO_2 ratio) being the only criterion that consistently stratified patients according to mortality risk.[191] Thus, direct comparison between studies may be difficult.

Girardis et al. found statistically significant reduced ICU-mortality in the modified intention-to-treat cohort of patients who were invasively mechanically ventilated at baseline and randomised to a conservative oxygenation strategy (absolute risk reduction 10.9 percentage, 95% CI 1.5-20.2, p = 0.02).[65] Among invasively mechanically ventilated patients at baseline in the HOT-ICU trial (n = 1,689), the point estimate for 90-day mortality favoured a higher oxygenation strategy (RR 1.07, 95% CI 0.95-1.21), but without significant interaction with treatment allocation. All patients included in the ICU-ROX trial were mechanically ventilated (either IMV or NIV) at baseline, and the trial was neutral overall for all outcomes.[83] In the O_2-ICU trial 295 of 400 in the intention-to-treat analysis were invasively mechanically ventilated at randomisation, but no subgroup results from this trial have yet been published. As previously mentioned, all patients in the HYPERS2S trial were also invasively mechanically ventilated, and the trial suggested a non-significant benefit of a lower oxygenation strategy.[66]

On a sidenote, routine supplemental oxygen may aggravate myocardial injury in patients with ST-elevation myocardial infarction if patients are not hypoxaemic,[192] and has not been proven to be beneficial in patients with myocardial infarction in general[165,193–195], nor in non-hypoxaemic stroke patients.[196] Similar suggestions have also been made in trauma patients where liberal oxygen supplementation and hyperoxaemia have been linked with increased mortality.[197–200] These findings further emphasize the need for more solid evidence on the optimum oxygenation strategies for acutely and critically ill patients.

5.2. Oxygenation targets and parameters

PaO_2 was chosen as the measure to target oxygenation in the HOT-ICU trial for the following reasons: the peripheral oxygen saturation (SpO_2) may deviate from the SaO_2 in a number of clinical circumstances (e.g. due to poor peripheral blood circulation);[201,202] and due to the sigmoid shape of the oxygen saturation curve, small changes in SpO_2 or SaO_2 values on the upper end of the scale (i.e. above 98%) could result in large, unrecognised changes in PaO_2. The PaO_2 would therefore allow for a more accurate oxygenation therapy. However, as it is currently not feasible to continuously measure PaO_2 in the ICU, all patients were continuously monitored with SpO_2 surveillance. The correlation of the two measures was used in practice to guide oxygenation between PaO_2 measurements. Unfortunately, due to pragmatic considerations, no SpO_2 measurements were registered.

As a built-in safety feature of the eCRF, an automatic e-mail notification was produced whenever a major protocol violation occurred (see section 3.1.9). This allowed us to support the participating sites in maintaining protocol adherence and investigate any potential safety issues related to the allocated oxygenation targets. However, this system relied on timely data registration, which was difficult to achieve. There were several reasons for this, lack of staff and high patient flow being the most common. Delayed data registration would thus reduce our capability to maintain protocol adherence and identify safety issues in a timely manner. To mitigate this issue, we were available 24/7/365 on both e-mail and telephone, and could thus be contacted at all times in case of any urgent issues. This feature was in many cases used by the sites, even during weekends, evenings, or at night. Despite challenges with timely data registration, excellent separation in terms of oxygenation parameters (PaO_2, FiO_2, and SaO_2) between the two groups for the vast majority of the intervention period was achieved. This emphasises the sites' overall abilities to follow the protocol (see section 4.1.3), and thus the potential to implement targeted oxygenation therapy into clinical practice. For both PaO_2 and SaO_2, only limited overlap in reported measurements (displayed as IQRs) were observed, and only in the second half of the intervention period (Figure 3 and Figure 7). This was most likely due to limited data as the number of patients admitted beyond 40 days with available oxygenation data was very limited (approximately 3% of patients randomised, see Table 6), thus resulting in large statistical uncertainty, which is evident from the expanding IQRs on the oxygenation graphs (Figure 5). Also, it is reasonable to assume that most patients will either die or improve clinically during their ICU admission. Thus, the need for supplemental oxygen will presumably diminish gradually, resulting in a lower FiO_2 to obtain the allocated oxygenation target, and ultimately no supplemental oxygen being needed (i.e. $FiO_2 = 0.21$).

The pre-defined oxygenation target of 8 kPa in the lower oxygenation group was not reached in the HOT-ICU trial, based on our available data, as the reported median

PaO_2 was 9.4 kPa (95% CI 8.9-10.2), and the median patient-mean of the registered 12-hour lowest values were 8.3 kPa (7.9-8.8) (Paper I, Appendix A – supplement Figure S1).[132] We also observed a marked right skewed distribution of the registered PaO_2 measurements in the lower group (Figure 4). However, it is important to note that the histograms in Figure 4 represent individual PaO_2 registrations, and are thus not representative for how the oxygenation targets were achieved on a patient level. Patients with longer lengths of ICU admission will disproportionally dominate the data as they provided more information. In the higher group, the pre-defined target was generally reached, with a median PaO_2 of 12.4 kPa (95% CI 11.6-13.2) and a more normal distribution of registered measurements (Figure 4). There could be several potential explanations for this finding. One may well reflect on clinicians' reluctance to target values below 8 kPa due to the potential risk of desaturation and detrimental hypoxaemia, whereas targeting in the vicinity below 12 kPa was perhaps not viewed as being as problematic. Also, some patients included in the estimation of the overall oxygenation parameters were withdrawn from the trial due to withdrawal of consent, but consented to further data registration. These patients would therefore not necessarily maintain their allocated target with high fidelity, leading to potential distortion of the estimates. Despite no supplemental oxygen, patients could still reach PaO_2 levels above their allocated target as initial gas exchange impairments improved during ICU care. We found a total of 11.6% of all registered PaO_2 measurements in the lower group were at an FiO_2 of 0.21 (Figure 6), and 9.3% of all registered PaO_2 measurements were at an FiO_2 of 0.21 and above 8 kPa. This would naturally increase the overall PaO_2 estimate for this group. For comparison, 1.8% of all registered PaO_2 measurements in the higher group were at an FiO_2 of 0.21, and only 0.6% of all PaO_2 registrations were at an FiO_2 of 0.21 and above 12 kPa. Conversely, patients were at times unable to achieve their allocated target despite an FiO_2 of 1.00 due to severe respiratory impairment. This was the case in 7.7% of all PaO_2 registrations in the higher group and 3.7% in the lower group. However, none of these cases (being above-target at FiO_2 0.21 or below-target at FiO_2 1.00) were considered protocol violations (see section 3.1.9).

When considering the relationship between the registered PaO_2 and SaO_2 measurements in the entire HOT-ICU cohort, it is apparent that this is subject to substantial uncertainty relating to both parameters; one measure of PaO_2 could correspond to a range of SaO_2 values, and vice versa (Figure 9). Also, PaO_2 values above 20 kPa (but in fact already from 8 kPa) resulted in SaO_2 values ≈ 99-100%, albeit with some spread. Up until approximately 10-12 kPa the relationship between the two was linear (or curve-linear) with increasing PaO_2 values resulting in increasing SaO_2 values (Figure 9). No sophisticated attempt to model the relationship was considered for this thesis. Besides inaccuracy in the measuring apparatus, the apparent uncertainty of correlation can be explained by the right or left shifting of the oxygen-haemoglobin dissociation curve caused by changes in the pH, partial pressure of carbon-dioxide, temperature, and concentration of 2,3-bisphosphoglyceric.[203–207] In turn, all of these parameters are influenced by the

individual patient's condition at the moment of sampling. Thus, in order to prevent hyperoxaemia, one should avoid the top-end of the oxygen saturation spectrum when targeting the oxygenation, but also frequently measure the PaO_2 as it may change in relation to the patient's clinical condition.

As displayed in Figure 10, the PaO_2/FiO_2 ratio improved remarkably similarly in both groups from the time of inclusion until approximately day 40, where the amount of data was very limited. Slightly higher PaO_2/FiO_2 ratios were observed in the lower group (not tested statistically). This could perhaps be explained by the inherent intricate relationship between the two parameters, where the PaO_2 increases disproportionally compared to the FiO_2 when the latter is increased.[42,43] The complex relationship between the two variables complicates the interpretation of the ratio, and may even be misleading (see section 1.3).[44]

We decided to report data on all oxygenation parameters for the entire intervention period for the sake of transparency, and due to the non-normal distribution of these variables we chose to graphically display the median values and IQRs. With data up to 90 days post-randomisation, the HOT-ICU trial is the trial with the longest intervention period and with most reported information on oxygenation parameters to date. In comparison, The OXYGEN-ICU trial reported the overall median time-weighted estimates of FiO_2 and PaO_2, with no information on daily measures, but only the distributions.[65] Similarly, only overall estimates on oxygenation parameters were provided by Martin et al. (TOXYC feasibility trial) and Yang et al. (POSDOT pilot trial).[161,162] Barrot et al. ($LOCO_2$ trial) reported time-weighted averages for PaO_2, FiO_2, and SaO_2 for the first 7 days after randomisation,[82] as did Panwar et al. (CLOSE I study) for SpO_2, PaO_2, and FiO_2 in their trial.[68] Gelissen et al. (O_2-ICU trial) reported the daily time-weighted median PaO_2 for the first 15 study days (day of admission + 14 intervention days),[84] whilst the investigators of the ICU-ROX trial reported daily time-weighted average PaO_2 and FiO_2 for the first 10 days, and the daily highest and lowest values for both parameters until day 28 after randomisation.[83]

Girardis et al. only reported oxygenation data from the modified-intention-to treat cohort (patients with actual length of ICU admission ≥72 hours), but failed to provide any data on SpO_2 even though their intervention was partly based on this parameter.[65] Also, the median FiO_2 in the 'conventional' group was 0.39 (IQR 0.35-0.42) and not the intended ≥0.40, albeit 25% of patients had a median FiO_2 >0.42. In the 'conservative' group the median FiO_2 was 0.36 (IQR 0.30-0.40), yielding a median absolute difference of only three percentage points. The apparently large impact on mortality reported in this trial is most likely due to chance findings, and not due to the interventions applied as the between-group difference in oxygenation is quite miniscule, albeit statistically significant. It is hard to imagine that a decrease in FiO_2 of 0.03 would result in an almost 50% relative reduction in ICU mortality. For comparison, Yang and Zhang reported an approximately 40% relative reduction in ICU mortality and difference in PaO_2 of only 0.8 kPa (estimated 95% CI 0.6-1.2 kPa)

between the two groups in their trial.[173] Again, this effect on ICU mortality is most likely also due to chance, as previously discussed. Such a findings have yet to be corroborated in other trials. In the HOT-ICU trial, the median difference in FiO_2 during the intervention period was 13 percentage points (see section 4.1.3). The achieved PaO_2 in the OXYGEN-ICU trial was 11.6 kPa (IQR 10.5-12.9) and 13.6 kPa (11.7-15.4) in the conservative and conventional groups, respectively, thus having substantial overlap.[65] No comparable data from the ICU-ROX trial has been published, though patients in this trial did spend more time at FiO_2 0.21 in the conservative group as compared to usual care.[83] However no overall estimate of PaO_2 during the intervention is provided, except graphs of daily time-weighted PaO_2 supplied in the supplement to the trial's main publication.[83] Oxygenation parameters are presented as means and standard errors, and only with error bars pointing away from the compared group, thus not illustrating any potential overlap. Such a depiction, could to the untrained eye, allude to a greater separation than actually achieved. Despite not achieving the target in the 'high-normal' group as discussed above, Gelissen et al. reported that the time-weighted PaO_2, SaO_2, SpO_2, and FiO_2 were significantly lower in the low-oxygen group as compared to the high-normal group.[84] In the HYPERS2S trial, oxygenation was significantly different between the two groups during the intervention period (24 hours after randomisation), but comparable at 72 hours post-randomisation.[66] However, the intervention period in this trial was only the first 24 hours of ICU admission after randomisation.

5.3. Strengths and limitations

This PhD thesis focusses on oxygenation strategies among critically ill adult patients admitted to the ICU, and combines the findings from the largest RCT on targeted oxygenation in ICU patients to date, including a pre-specified, secondary, Bayesian analysis of the trial, and the most recently updated Cochrane review on higher versus lower oxygenation strategies in adult ICU patients The methodology applied across all investigations has been selected with great care, rigour, and implemented to the highest scientific level possible. However, the findings presented in this thesis are not necessarily transferable to other patients categories than those investigated, e.g. acutely ill patients outside the ICU, and non-critically ill patients.

In addition to presenting data on almost three times as many patients as the second largest, the ICU-ROX trial (n = 1,000),[83] the HOT-ICU trial is also the first to be powered at 90% to demonstrate a mortality difference, and designed to demonstrate the smallest absolute mortality difference yet of five percentage points. Also, the HOT-ICU trial was completed with inclusion of the planned number of patients, with only minimal loss to follow-up, and with a clear separation of oxygenation between the two groups as discussed above. Additionally, both the trial

protocol and statistical analysis plan were published prior to inclusion of the last patient.[123,124] Few changes to the trial protocol were made during trial conduct, and all pertained to updates of investigators. Details on changes to the protocol are presented elsewere.[132] All pre-defined short-term outcomes were reported on, and reported as pre-defined, except for a minor change in the reporting of the secondary analysis of the primary outcome (see section 3.1.8).

To assess the robustness of the intervention effect, the conduct of sensitivity analyses (or per-protocol analyses) is essential, and such are optimally pre-specified,[208] as in the case of the HOT-ICU trial.[123,124] This is so, since the effect of protocol adherence may affect the results from the primary intention-to-treat analysis (where patients remain in their designated groups despite deviations), and thus the overall interpretation of a trial's results.[209] We are currently working on the previously mentioned per-protocol analyses (see section 3.1.9) in addition to a range of post-hoc defined analyses, and are therefore unable to present the results at this moment.

Due to the nature of oxygen therapy in the ICU, the HOT-ICU trial was not blinded to clinicians, patients, or relatives, but assessment of the primary outcome, mortality, is not very likely to be influenced by knowledge of target allocation. Though inability to blind an intervention does not necessarily result in a high risk of bias, according to the updated RoB-2 evaluation tool,[128] inadequate blinding may still introduce risks of underestimating SAEs and overestimating positive intervention effects.[210,211] In an attempt to mitigate this issue, all statistical analyses were conducted in a blinded manner, and two abstracts were written prior to unblinding the results (presented in the supplement to Paper I - Appendix A).[132] We saw potentially different intervention effects when considering participating sites, as per a post-hoc analysis. However, randomisation was stratified for site, and we also adjusted for this effect in the analyses.

Due to the pragmatic design of the HOT-ICU trial, and in an attempt to minimise the burden of data registration on the participating sites, only the highest and lowest PaO_2 measurements in pre-defined 12-hour intervals were registered, despite additional daily measurements during ICU admission. On average, patients had six arterial blood gas samples (ABGs) taken per day, with most conducted during the beginning of the ICU admission and fewer (or none) at the end, as displayed in Figure 12. This was less than previously reported in the Danish observational cohort study, which was used as a foundation for the HOT-ICU trial design,[55] however a similar number of daily ABGs was reported in other studies.[28,54] This approach did not allow for calculation of time-weighted estimates, as has been reported in other trials,[65,68,84,161,162] but provided an estimate of the oxygenation range in which the patients had spent their time. By only registering the extremes of the oxygenation parameters, the higher values could potentially have a greater influence on the estimates than the lower, since there is a large range of values above the designated

targets, but only a limited range below. This would potentially diminish the separation of the two groups, as extreme peak measurements would exaggerate the overall estimates. Also, as only the concomitant measurements of SaO_2 and FiO_2 were registered, values of these two parameters would not necessarily represent the highest and lowest in any given interval. We are therefore not able to, with absolute certainty, account for the exact details of these variables. Moreover, most of the data on oxygenation was collected within the first two to three weeks of intervention, consequently making inferences on the effects of long-term targeted oxygen therapy more difficult. We did not register ethnicity, and are therefore unable to evaluate the potential influence of skin colour, as SpO_2 measurement may differ with skin colour.[212] Time needed to complete recruitment was longer than first anticipated, as the trial was initially planned to be completed within two years. A long inclusion period could potentially lead to a drift in treatment fidelity, as a consequence of familiarity with the trial protocol and clinicians/nurses perceptions of the intervention effect (be that equipoise or belief in the superiority of one of the oxygenation targets). Due to the inherent difficulty with estimating FiO_2 in open oxygen supplementation systems, we used standardised conversion tables, knowing that this would probably lead to overestimation of the 'true' FiO_2 and consequently also misestimating the PaO_2/FiO_2 ratio.

The HOT-ICU trial's pragmatic, multicentre design (35 ICUs in 7 countries) allows for a high level of external validity and thus applicability to ICUs in general. All treatments prescribed to patients, other than adjustment of the FiO_2 in order to reach the allocated oxygenation target, were at the discretion of the treating physician, and thus reflected standard clinical practice at the involved sites. The applied oxygenation targets in the trial were chosen based on both a retrospective, multicentre, observational cohort study and a questionnaire among Northern European ICU physicians.[34,55] Also, at the time of design the recommended oxygenation target for patients with ARDS was 7.3-10.7 kPa based on trials by the ARDS-Network.[37,174] However, the oxygenation targets applied in the HOT-ICU trial may not be comparable to standard-of-care in other ICUs, thus limiting generalisability. This could however also be the case for what was defined as 'usual care' in the other RCTs investigating this matter. In addition, we restricted inclusion to patients acutely admitted to the ICU with acute hypoxaemic respiratory failure, resulting in baseline PaO_2/FiO_2 ratios corresponding to moderate-to-severe ARDS. Most patients were admitted due to acute medical conditions (most often pneumonia). Therefore, generalisability to other patient categories, e.g. those with less severe degrees of pulmonary insufficiency (i.e. higher PaO_2/FiO_2 ratios), or lower baseline risks of death may be limited.

As was evident, the mortality in the trial was higher than anticipated and this could be due to several reasons. The composition of admissions was different to that expected, with more than 80% of medical admissions in our trial and around 40% in the initial cohort study on which power calculation was partly based.[55] As we

intended to study the effects of targeted oxygenation therapy in patients with hypoxaemic respiratory failure, the inclusion criteria were design such that we anticipated all patients to have a PaO_2/FiO_2 ratio below 40 kPa at baseline, corresponding to the cut-off for mild ARDS. This was the case for 96.8% of all included patients, but more than 88% of all patients had ratios within the range of moderate-to-severe ARDS at inclusion. The overall mortality in the HOT-ICU trial also corresponded to that of patients with moderate-to-severe ARDS.[213] Yet, only 12.8% were diagnosed with ARDS at baseline by clinicians. This group is most likely not correctly identified and probably underestimated. Again, due to practical concerns, we did not register if patients had chest x-rays taken, timing of the respiratory insults, nor origin of any pulmonary oedema, and are therefore not able to validate the ARDS diagnoses post-hoc, nor to estimate the number of patients who developed ARDS after randomisation. Poor clinical recognition of ARDS is not uncommon, despite the high mortality risk for patients with such a diagnosis.[213] However development of ARDS may be somewhat delayed in relation to development of hypoxaemic respiratory failure,[213] and may thus partly explain the apparent discrepancy between baseline PaO_2/FiO_2 ratios and diagnosis of ARDS in the HOT-ICU trial.

The HOT-ICU trial was conducted a part of an international collaboration: the Centre for Research in Intensive Care (CRIC, cric.nu), which is built on a highly successful and cost-effective clinical research model that has produced several large-scale international, multicentre RCTs.[105,214–216] Three additional RCTs based on the CRIC-collaboration are currently enrolling,[217–219] and one has recently completed enrolment.[220]

All the strengths and limitations of the HOT-ICU trial are relevant to the Bayesian analysis of the trial. However, of particular additional strength to this study are the following: the protocol for this study was published prior to randomisation of the last patient in the HOT-ICU trial thus avoiding a data driven post-hoc design;[96] the results of this study were consistent across several different pre-specified priors (minimally informative, sceptic, and evidence-based); and the presence of HTE was evaluated using two difference approaches. Choice of baseline variables for the HTE analyses were made after the start of the HOT-ICU trial, but before completion of recruitment. Therefore, this decision was partly based on availability. Even though the SOFA score was not designed as a predictor of mortality, it has been demonstrated to predict mortality reasonably well,[221–223] though not with the same discriminatory power as, for example, the SAPS-II and APACHE-IV scores (both dedicated mortality predictions scores).[222,223] Also, when estimating the SOFA score, there are multiple ways for patients to obtain the same score which may differ in their effects on outcomes (e.g. risk of dying). Therefore, the relationship between increments in baseline SOFA score and mortality is not necessarily linear. A dedicated mortality prediction score would have been preferred, but was not available. Another limitation is the fact that we only estimated the baseline SOFA score, as

serial evaluation of the SOFA score can be a good predictor of mortality.[224] As discussed previously, the PaO_2/FiO_2 ratio is highly sensitive to the denominator, and thus does not necessarily have a monotonic relationship with the level of hypoxaemic respiratory failure, and hence with mortality.[42,43] Choice of baseline variables are discussed in more detail in the study protocol and Paper II (Appendix B).[96,133] Also, in this study we only incorporated a linear relationship on the log-OR scale when investigating HTE on the continuous scale, even though other models could have been chosen (and were discussed during the design of the study), e.g. a quadratic or cubic relationship, or using other smoothing functions. We opted not to do this, both for the sake of simplicity, but also to limit the number of analyses, and at the same time to limit the risk of overfitting and producing spurious results due to high model flexibility. The logistic regression model, with a log-linear relationship, results in a sigmoid relationship on the natural scale, and hence inherently allows for some non-linearity of the variable.

An additional advantage of the Bayesian approach is the potential to incorporate different assumptions of effects within the model (i.e. priors). An example of this is the post-hoc, Bayesian Analysis published by Goligher et al. in 2018[100] of the 'ECMO to Rescue Lung Injury in Severe ARDS' (EOLIA) trial.[99] The original trial was stopped early due to futility. It failed to demonstrate a statistically significant reduction in 60-day mortality by using early ECMO as compared to standard care for severe hypoxaemic respiratory failure; RR in favour of early ECMO 0.79 (95% CI 0.55-1.04, p = 0.09).[99] The authors of the Bayesian analysis constructed a range of priors (minimally informative, strongly enthusiastic, moderately enthusiastic, sceptical, and strongly sceptical) for the mortality effect. Not surprisingly, the more enthusiastic the defined priors (i.e. in favour of the intervention) the more likely an intervention effect was to favour early ECMO. As such, priors may greatly influence the posterior distribution and it is essential to pre-define all priors in order to mitigate spurious findings, thus emphasising the subjectivity of the Bayesian approach.[93] To avoid any issues with post-hoc design of priors in the HOT-ICU trial, we used three pre-defined priors, of which two were used for the sensitivity analysis purposes.[96] One was a *sceptic* prior, centred on no effect (mean OR 1.00) and with 95% probability between 0.75 and 1.34. This prior would thus favour estimates closer to no effect and was generally sceptical of large effect sizes, given the narrower probability interval. The other was an *evidence based* prior constructed from an updated Cochrane meta-analysis,[81] with a slight inclination to favour the lower oxygenation target, as the mean OR was 0.93 and with 95% probability between 0.72-1.20. Even though the evidence-based prior favoured the lower target, it was even more sceptical of large effect sizes than the sceptic prior. Details on priors are provided in the study protocol and Paper II (Appendix B).[133] Both sensitivity analyses were consistent with the results of the primary Bayesian analysis.[96,133] Again, this was not surprising as the amount of data from the HOT-ICU trial was expected to dominate the posterior distributions, and this minimised any effects of the priors. Unfortunately, as no other

trial on lower versus higher oxygenation strategies has been analysed in a Bayesian framework, direct comparison of such is not possible in this context.

When considering the updated Cochrane review, the methodology applied was both rigorous and up-to-date, thus affirming the reliability of the findings. Several post-hoc changes were made from the protocol to the updated version of the review, e.g. changes to subgroup analyses, focus on outcomes at maximum follow-up only, and expansion of search strategy. All changes are listed in the section 'Differences between protocol and review' in Paper III (Appendix C).[134] All meta-analyses were perform within the RevMan Web domain, and the choice of meta-analysis models was thus restricted, despite an alternative random-effects models with proposed smaller risk of false-positive findings exist.[225] We reported the most conservative estimate, being the model with the point estimate closest to the null-effect. Also, by multiplicity-adjusting the significance levels and by conducting Trial Sequential Analyses the problem of increased error in the random-effects model rate was to some extent lessened. The trials identified in the updated Cochrane review were highly diverse, both in terms of inclusion/exclusion criteria, and thus the studied populations (as previously discussed), but in particular when considering the applied interventions. The trials did not define the lower and higher oxygenation strategies similarly, nor were the means to implement such strategies the same; some used oxygenation targets based on measurements of SpO_2, PaO_2, or SaO_2 (or combinations) whereas others used fixed levels of FiO_2 (see Table 9). The duration of the applied interventions also varied immensely (from a few hours up to 90 days). This is of great importance, as both the level of oxygenation but also the time spent at hyperoxaemic levels have previously been associated with increased mortality.[28] For further elaboration of the clinical diversities in the identified trials see Paper III (Appendix C).[134] The overall certainty of evidence for all outcomes was low or very low, due to risks of bias, limited data, differences in investigated populations, and large heterogeneity of applied interventions in the identified trials. All of these factors, and in particular the latter, impede the overall interpretation of the meta-analyses. A new tool for systematically appraising the clinical diversity in RCTs has recently been suggested, thus allowing adjustment for relevant factors by e.g. meta-regression.[226] In this context, clinical diversity is considered within four domains (setting, population, intervention, and with 11 overall items each assigning 0-2 points for the level of diversity [0 = 'low', 1 = 'moderate' or 'unknown' and 2 = 'high']), which could plausibly influence the overall effect estimate. Such an investigation could perhaps shed more light on this matter, e.g. by means of meta-regression analyses.

6. CONCLUSION AND PERSPECTIVES

In summary, the findings presented in this thesis add to the combined sum of knowledge concerning targeted oxygenation therapy in the adult ICU patient.

Though the completion of the HOT-ICU trial has provided more data on this matter than previously published, no firm conclusion regarding the benefits or harms of higher versus lower oxygenation strategies could be drawn, and we are still uncertain about the true effect of targeted oxygenation therapy. However, based on the evidence presented in this thesis, large effects on mortality or on the risk of serious adverse events are not likely when targeting oxygenation within the 'normal range'. Currently, we cannot proclaim a 'one size fits all' model as best, but need to recall that oxygen is still classified as a medical drug, and should thus continue to be prescribed with both the benefits of avoiding hypoxaemia and the harms from hyperoxaemia in mind. Oxygen should correspondingly only be prescribed to those patients needing it (and at the lowest safe amount), and not as a 'safety precaution' to all patients, as this will most likely only provide a 'false safety' and potentially induce harm due to the risks associated with hyperoxaemia.

In situations of limited resources and restricted oxygen delivery, it could be appropriate to implement a more restrictive oxygen regimen. This has for instance been very relevant in many countries around the world during the COVID-19 pandemic with numerous hospitals being unable to obtain the required volumes of oxygen due to global shortages, logistical breakdowns, and the sudden increase in the use of oxygen.

A pragmatic approach to implementing the findings presented in this thesis could be to avoid an oxygen saturation (SaO_2 or SpO_2) of 99-100%. At this level it is impossible to guard against hyperoxaemia, as the saturation most often remains unchanged despite large changes in PaO_2, due to the nature of the oxygen-haemoglobin oxygen dissociation curve. This also puts the clinician in a position where detection of a deterioration in the patient's clinical state can be delayed, as the PaO_2 needs to drop substantially before being clearly displayed by means of SaO_2 or SpO_2. If one targets the oxygenation using frequent PaO_2 measurements and aims for 8-10 kPa instead (using SpO_2 as a guide between PaO_2 measurements), this will often result in a saturation on the steeper part of the oxygen-haemoglobin dissociation curve. Clinicians would thus be warned earlier should the patient's oxygenation drop, and allow timely implemented interventions to correct such a trajectory.

In the updated Cochrane review we found indications of decreased risk of developing sepsis during ICU admission if a lower oxygenation strategy was employed, albeit with very low certainty of evidence. Conversely, results from the HOT-ICU trial

suggested potential harm (i.e. increased mortality) from a lower oxygenation target among patients admitted to the ICU with both hypoxaemic respiratory and circulatory failure (i.e. shock). As development of sepsis, and in particular septic shock, is highly correlated with increased risk of death, these findings warrant further investigations.

Further investigations on the optimum oxygenation strategy in critically ill patients with e.g. COPD, traumatic brain injury, ARDS, and those resuscitated from cardiac arrest is also highly clinically relevant, as current knowledge is insufficient for these specific patient categories.

Future trials should be designed in order to minimise sources of bias, and ensure adequate separation between the applied oxygenation strategies. They should also be sufficiently powered in order to evaluate not only hard endpoint (e.g. mortality), but also to consider patient-centred, and other clinically relevant outcomes, e.g. quality-of-life, or long-term pulmonary and cognitive impairments. In the HOT-ICU trial we are currently investigating such effects, as all patients surviving to one year have been evaluated in terms of self-reported health-related quality-of-life. Additionally, Danish patients at selected sites have been invited to participate in extensive tests of their pulmonary and cognitive functions one year after randomisation. The findings of these investigations will undoubtably further our understanding of the long-term effects of targeted oxygenation therapy in the ICU, and contribute to the advance of clinical practice for the benefit of patients.

REFERENCES

1. Costa KM, Accorsi-Mendonça D, Moraes DJA, et al. Evolution and physiology of neural oxygen sensing. *Frontiers in Physiology*. 2014;5(August):1–16. doi:10.3389/fphys.2014.00302

2. Priestley J. An account of further discoveries in air. *Philosophical Transactions of the Royal Society of London*. 1775;65:384–94. doi:10.1098/rstl.1775.0039

3. Scheele CW. Chemische Abhandlung von der Luft und dem Feuer. Upsala, Sweden & Leipzig, Germany: Verlegt von Magn. Swederus Buchhandler, zu finden bey S. L. Crusius, 1777.

4. Lavoisier A-L de. Memoir on the Nature of the Principle Which Combines with Metals during Their Calcination and Increases Their Weight. Mémoires de l'Académie Royale des Sciences, 1778.

5. Lavoisier A-L de. Traité élémentaire de chimie , présenté dans un ordre nouveau, et d'après les découvertes modernes. Paris: Cuchet, Libraire, 1789.

6. Harvey W. Exercitatio Anatomica de Motu Cordis et Sanguinis in Animalibus. Frankfurt: William Fitzer, 1628.

7. Bert P. La Pression Barométrique: Recherches de Physiologie Expérimentale. Paris: Lbiraire de l'Académie de Médecine, G. Masson, 1878.

8. Smith JL. The Pathological Effects Due To Increase of Oxygen Tension in the Air Breathed. *Journal of Physiology*. 1899;24(1):19–35. doi:10.1113/jphysiol.1899.sp000746

9. Beasley R, Chien J, Douglas J, et al. Thoracic Society of Australia and New Zealand oxygen guidelines for acute oxygen use in adults: Swimming between the flags. *Respirology*. 2015;20(8):1182–91. doi:10.1111/resp.12620

10. O'Driscoll BR, Howard LS, Earis J, et al. BTS guideline for oxygen use in adults in healthcare and emergency settings. *Thorax*. 2017;72(6):i1-90. doi:10.1136/thoraxjnl-2016-209729

11. The Danish Health Authority. National Klinisk Retningslinje for iltbehandling til den akut syge voksne patient [Internet]. Available from: https://www.sst.dk/da/udgivelser/2019/nkr-iltbehandling-til-den-akut-

syge-voksne-patient

12. Beasley R, Chien J, Douglas J, et al. Target oxygen saturation range: 92–96% Versus 94–98%. *Respirology*. 2017;22(1):200–2. doi:10.1111/resp.12879

13. Fan E, Sorbo L Del, Goligher EC, et al. Guidelines of the American Thoracic Society (ATS), European Society of Intensive Care Medicine (ESICM), and Society of Critical Care Medicine (SCCM). *American Thoracic Society Documents*. 2017;195(9):1253–63. doi:10.1164/rccm.201703-0548ST

14. Leigh JM. Early treatment with oxygen. *Anaesthesia*. 1974;29(2):194–208. doi:10.1111/j.1365-2044.1974.tb00623.x

15. Zhu H, Santo A, Trush M, et al. Oxygen and Oxygen Toxicity: The Birth of Concepts. *Reactive Oxygen Species (Apex)*. 2016;1(1):1–8. doi:10.20455/ros.2016.801

16. Helmerhorst HJF, Schultz MJ, van der Voort PHJ, et al. Bench-to-bedside review: The effects of hyperoxia during critical illness. *Critical Care*. 2015;19(1):1–12. doi:10.1186/s13054-015-0996-4

17. Clark JM, Lambertsen CJ. Pulmonary oxygen toxicity: a review. *Pharmacological Reviews*. 1971;23(2):37–133.

18. Jakcson RM. Pulmonary Oxygen Toxicity. *Chest*. 1985;88(6):900–5.

19. Chow C-W, Teresa Herrera Abreu M, Suzuki T, et al. Oxidative Stress and Acute Lung Injury. *Am J Respir Cell Mol Biol*. 2003;29:427–31. doi:10.1165/rcmb.F278

20. Singer M, Young PJ, Laffey JG, et al. Dangers of hyperoxia. *Critical Care*. 2021;25:1–15. doi:10.1186/s13054-021-03815-y

21. 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

22. Donald KW. Oxygen poisoning in man. *British Medical Journal*. 1947;1(4507):712–7. doi:10.1136/bmj.1.4507.712

23. Barber RE, Lee J, Hamilton WK. Oxygen Toxicity in Man. *New England Journal of Medicine*. 1970;283(27):1478–84. doi:10.1056/nejm197012312832702

24. Rothen HU, Sporre B, Engberg G, et al. Prevention of atelectasis during

general anaesthesia. *The Lancet*. 1995;345(8962):1387–91. doi:10.1016/S0140-6736(95)92595-3

25. Tipple TE, Ambalavanan N. Oxygen Toxicity in the Neonate: Thinking Beyond the Balance. *Clinics in Perinatology*. 2019;46(3):435–47. doi:10.1016/j.clp.2019.05.001

26. Floyd TF, Clark JM, Gelfand R, et al. Independent cerebral vasoconstrictive effects of hyperoxia and accompanying arterial hypocapnia at 1 ATA. *Journal of Applied Physiology*. 2003;95(6):2453–61. doi:10.1152/japplphysiol.00303.2003

27. Bak Z, Sjöberg F, Rousseau A, et al. Human cardiovascular dose-response to supplemental oxygen. *Acta Physiologica*. 2007;191(1):15–24. doi:10.1111/j.1748-1716.2007.01710.x

28. Helmerhorst HJF, Arts DL, Schultz MJ, et al. Metrics of arterial hyperoxia and associated outcomes in critical care. *Critical Care Medicine*. 2017;45(2):187–95. doi:10.1097/CCM.0000000000002084

29. Deichmann WB, Henschler D, Holmstedt B, et al. What is there that is not poison? A study of the Third Defense by Paracelsus. *Archives of Toxicology*. 1986;58(4):207–13. doi:10.1007/BF00297107

30. Barbateskovic M, Schjørring OL, Krauss SR, et al. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (Review). *Cochrane Database of Systematic Reviews*. 2019;(11):Art. No.: CD012631. doi:10.1002/14651858.CD012631.pub2

31. You J, Fan X, Bi X, et al. Association between arterial hyperoxia and mortality in critically ill patients: A systematic review and meta-analysis. *Journal of Critical Care*. 2018;47(2018):260–8. doi:10.1016/j.jcrc.2018.07.014

32. Li X, Liu D, Liu C, et al. Conservative versus liberal oxygen therapy in relation to all-cause mortality among patients in the intensive care unit: A systematic review of randomized controlled trials with meta-analysis and trial sequential analysis. *Medicina Intensiva*. 2021. doi: 10.1016/j.medin.2021.08.006 doi:10.1016/j.medin.2021.08.006

33. Li L, Zhang Y, Wang P, et al. Conservative versus liberal oxygen therapy for acutely ill medical patients: A systematic review and meta-analysis. *International Journal of Nursing Studies*. 2021;118:103924. doi:10.1016/j.ijnurstu.2021.103924

34. Schjørring OL, Toft-Petersen AP, Kusk KH, et al. Intensive care doctors' preferences for arterial oxygen tension levels in mechanically ventilated patients. *Acta Anaesthesiologica Scandinavica*. 2018;62(10):1443–51. doi:10.1111/aas.13171

35. Abdelsalam M. Permissive Hypoxemia. *Chest*. 2006;129(1):210–1. doi:10.1378/chest.129.1.210

36. Martin DS, Grocott MPW. Oxygen therapy in critical illness: Precise control of arterial oxygenation and permissive hypoxemia. *Critical Care Medicine*. 2013;41(2):423–32. doi:10.1097/CCM.0b013e31826a44f6

37. The 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. doi:10.1097/00132586-200102000-00017

38. Gilbert-Kawai ET, Mitchell K, Martin D, et al. Permissive hypoxaemia versus normoxaemia for mechanically ventilated critically ill patients. *Cochrane Database of Systematic Reviews*. 2014;May 7(5):CD009931. doi:10.1002/14651858.CD009931.pub2

39. Pratter MR, Irwin RS. Extrapulmonary Causes of Respiratory Failure. *Journal of Intensive Care Medicine*. 1986;1(4):197–217. doi:10.1177/088506668600100405

40. Pelosi P, D'Onofrio D, Chiumello D, et al. Pulmonary and extrapulmonary acute respiratory distress syndrome are different. *European Respiratory Journal*. 2003;22(Suppl. 42):48–56. doi:10.1183/09031936.03.00420803

41. Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: The Berlin definition. *JAMA - Journal of the American Medical Association*. 2012;307(23):2526–33. doi:10.1001/jama.2012.5669

42. Karbing DS, Kjærgaard S, Smith BW, et al. Variation in the PaO2/FiO2 ratio with FiO2: Mathematical and experimental description, and clinical relevance. *Critical Care*. 2007;11(6):1–8. doi:10.1186/cc6174

43. Aboab J, Louis B, Jonson B, et al. Relation between PaO2/FiO2 ratio and FiO2: a mathematical description. *Intensive Care Medicine*. 2006;32(10):1494–7. doi:10.1007/978-3-642-28270-6_15

44. Broccard A. Making sense of the PaO2/FIO2 ratio in patients with acute respiratory distress syndrome. *OA Critical Care*. 2013;June 01(9):1–16.

45. Feiner JR, Weiskopf RB. Evaluating Pulmonary Function: An Assessment of PaO2/FiO2. *Critical Care Medicine*. 2017;45(1):e40–8. doi:10.1097/CCM.0000000000002017

46. de Jonge E, Peelen L, Keijzers PJ, et al. Association between administered oxygen, arterial partial oxygen pressure and mortality in mechanically ventilated intensive care unit patients. *Critical Care*. 2008;12(6):1–8. doi:10.1186/cc7150

47. De Graaff AE, Dongelmans DA, Binnekade JM, et al. 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. doi:10.1007/s00134-010-2025-z

48. Eastwood G, Bellomo R, Bailey M, et al. Arterial oxygen tension and mortality in mechanically ventilated patients. *Intensive Care Medicine*. 2012;38(1):91–8. doi:10.1007/s00134-011-2419-6

49. Panwar P, Capellier G, Schmutz N, et al. Current oxygenation practice in ventilated patients-an observational cohort study. *Anaesth Intensive Care*. 2013;41:505–5014.

50. Dahl RM, Grønlykke L, Haase N, et al. Variability in targeted arterial oxygenation levels in patients with severe sepsis or septic shock. *Acta Anaesthesiologica Scandinavica*. 2015;59(7):859–69. doi:10.1111/aas.12528

51. 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(October):1–7. doi:10.1038/srep35133

52. Dennis D, Torre L, Baker S, et al. A snapshot of the oxygenation of mechanically ventilated patients in one Australian intensive care unit. *Anaesthesia and Intensive Care*. 2017;45(3):359–64.

53. Egi M, Kataoka J, Ito T, et al. Oxygen management in mechanically ventilated patients: A multicenter prospective observational study. *Journal of Critical Care*. 2018;46:1–5. doi:10.1016/j.jcrc.2018.03.024

54. Kraft F, Andel H, Gamper J, et al. Incidence of hyperoxia and related in-hospital mortality in critically ill patients: a retrospective data analysis. *Acta Anaesthesiologica Scandinavica*. 2018;62(3):347–56. doi:10.1111/aas.13047

55. Schjørring OL, Jensen AKG, Nielsen CG, et al. Arterial oxygen tensions in mechanically ventilated ICU patients and mortality: a retrospective,

multicentre, observational cohort study. *British Journal of Anaesthesia*. 2020;124(December):1–10. doi:10.1016/j.bja.2019.12.039

56. Aggarwal NR, Brower RG, Hager DN, et al. Oxygen Exposure Resulting in Arterial Oxygen Tensions Above the Protocol Goal Was Associated With Worse Clinical Outcomes in Acute Respiratory Distress Syndrome. *Critical care medicine*. 2018;46(4):517–24. doi:10.1097/CCM.0000000000002886

57. Ruggiu M, Aissaoui N, Nael J, et al. Hyperoxia effects on intensive care unit mortality: a retrospective pragmatic cohort study. *Critical Care*. 2018;22(1):4–6. doi:10.1186/s13054-018-2142-6

58. Palmer E, Post B, Klapaukh R, et al. The association between supraphysiologic arterial oxygen levels and mortality in critically ill patients a multicenter observational cohort study. *American Journal of Respiratory and Critical Care Medicine*. 2019;200(11):1373–80. doi:10.1164/rccm.201904-0849OC

59. Harvey J, Jayawardena D, Ramanan M. Oxygen exposure as quantified by time-weighted area under curve for arterial oxygen content is associated with mortality in mechanically ventilated critically ill patients. *Research and Opinion in Anesthesia and Intensive Care*. 2020;7:197–204. doi:10.4103/roaic.roaic_84_19

60. van den Boom W, Hoy M, Sankaran J, et al. The Search for Optimal Oxygen Saturation Targets in Critically Ill Patients: Observational Data From Large ICU Databases. *Chest*. 2020;157(3):566–73. doi:10.1016/j.chest.2019.09.015

61. Zhou D, Li Z, Shi G, et al. Time spent in oxygen saturation 95-99% is associated with reduced mortality in critically ill patients with mechanical ventilation. *Critical Care*. 2020;24(414):1–4. doi:10.1186/s13054-020-03126-8

62. Rachmale S, Li G, Wilson G, et al. Practice of Excessive FIO2 and Effect on Pulmonary Outcomes in Mechanically Ventilated Patients With Acute Lung Injury. *Respiratory Care*. 2012;57(11):1887–93. doi:10.4187/respcare.01696

63. Ramanan M, Fisher N. The Association between Arterial Oxygen Tension, Hemoglobin Concentration, and Mortality in Mechanically Ventilated Critically Ill Patients. *Indian Journal of Critical Care Medicine*. 2018;22(7):477–484. doi:10.4103/ijccm.IJCCM_66_18: 10.4103/ijccm.IJCCM_66_18

64. Madotto F, Rezoagli E, Pham T, et al. Hyperoxemia and excess oxygen use in early acute respiratory distress syndrome: insights from the LUNG SAFE study. *Critical Care*. 2020;24(1):1–17. doi:10.1186/s13054-020-2826-6

65. Girardis M, Busani S, Damiani E, 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 - Journal of the American Medical Association*. 2016;316(15):1583–9. doi:10.1001/jama.2016.11993

66. Asfar P, Schortgen F, Boisramé-Helms J, et al. Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S): a two-by-two factorial, multicentre, randomised, clinical trial. *The Lancet Respiratory Medicine*. 2017;5(3):180–90. doi:10.1016/S2213-2600(17)30046-2

67. Gomersall CD, Joynt GM, Sa FFA, et al. Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: A randomised, controlled pilot study. *Critical Care Medicine*. 2002;30(1):113–6.

68. Panwar R, Hardie M, Bellomo R, 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. doi:10.1164/rccm.201505-1019OC

69. Lång M, Skrifvars MB, Siironen J, et al. A pilot study of hyperoxemia on neurological injury, inflammation and oxidative stress. *Acta Anaesthesiologica Scandinavica*. 2018;62(6):801–10. doi:10.1111/aas.13093

70. Jakkula P, Care I, Reinikainen M, 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. doi:10.1007/s00134-018-5453-9

71. Young PJ, Mackle DM, Bailey MJ, 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.

72. Mazdeh M, Taher A, Torabian S, et al. Effects of normobaric hyperoxia in severe acute stroke: A randomized controlled clinical trial study. *Acta Medica Iranica*. 2015;53(11):676–80.

73. Taher A, Pilehvari Z, Poorolajal J, et al. Effects of normobaric hyperoxia in severe acute stroke: A randomized controlled clinical trial study. *Trauma Monthly*. 2016;21(1):1–5. doi:10.5812/traumamon.26772

74. Ishii K, Morimatsu H, Hyodo T, et al. Relationship between inspired oxygen concentration and atelectasis formation efter extubation. *Critical Care Medicine*. 2018;46(1 Suppl 1):533.

doi:10.1097/01.ccm.0000529104.66235.9e

75. Suzuki S, Eastwood GM, Glassford NJ, et al. Conservative oxygen therapy in mechanically ventilated patients: A pilot before-and-after trial. *Critical Care Medicine*. 2014;42(6):1414–22. doi:10.1097/CCM.0000000000000219

76. Helmerhorst HJF, Schultz MJ, Van Der Voort PHJ, et al. Effectiveness and Clinical Outcomes of a Two-Step Implementation of Conservative Oxygenation Targets in Critically Ill Patients: A before and after Trial. *Critical Care Medicine*. 2016;44(3):554–63. doi:10.1097/CCM.0000000000001461

77. Suzuki S, Eastwood GM, Goodwin MD, et al. Atelectasis and mechanical ventilation mode during conservative oxygen therapy: A before-and-after study. *Journal of Critical Care*. 2015;30(6):1232–7. doi:10.1016/j.jcrc.2015.07.033

78. Eastwood GM, Chan MJ, Peck L, et al. Conservative versus conventional oxygen therapy for cardiac surgical patients: A before-and-after study. *Anaesthesia and Intensive Care*. 2019;47(2):175–82. doi:10.1177/0310057X19838753

79. Chu DK, Kim LHY, Young PJ, et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA): a systematic review and meta-analysis. *The Lancet*. 2018;391(10131):1693–705. doi:10.1016/S0140-6736(18)30479-3

80. Siemieniuk RAC, Chu DK, Kim LHY, et al. Oxygen therapy for acutely ill medical patients: A clinical practice guideline. *British Medical Journal*. 2018;363(k4169):1–10. doi:10.1136/bmj.k4169

81. Barbateskovic M, Schjørring OL, Krauss SR, et al. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (Review). *Cochrane Database of Systematic Reviews*. 2019. doi: 10.1002/14651858.CD012631.pub2 doi:10.1002/14651858.CD012631.pub2

82. Barrot L, Asfar P, Mauny F, et al. Liberal or Conservative Oxygen Therapy for Acute Respiratory Distress Syndrome. *New England Journal of Medicine*. 2020;382(11):999. doi:10.1056/NEJMoa1916431

83. The ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Conservative Oxygen Therapy during Mechanical Ventilation in the ICU. *New England Journal of Medicine*. 2020;382(11):989–98. doi:10.1056/NEJMoa1903297

84. Gelissen H, Harm-Jan de Grooth, Smulders Y, et al. Effect of Low-Normal vs High-Normal Oxygenation Targets on Organ Dysfunction in Critically Ill Patients - A Randomized Clinical Trial. *JAMA - Journal of the American Medical Association*. 2021;August:1–9. doi:10.1001/jama.2021.13011

85. Jakobsen JC, Gluud C, Winkel P, et al. 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(1):1–12. doi:10.1186/1471-2288-14-34

86. Fisher RA. Statistical Methods for Research Workers. Edinburgh & Rothamsted: Oliver and Boyd, 1925.

87. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019;567:305–7. doi:10.1038/d41586-019-00857-9

88. Wasserstein RL, Schirm AL, Lazar NA. Moving to a World Beyond "p < 0.05." *American Statistician*. 2019;73(sup1):1–19. doi:10.1080/00031305.2019.1583913

89. Bayes T. An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London*. 1763;53:370–418.

90. Laplace P-S. Essai Philosophique sur les Probabilités. Paris: Mme. Ve. Courcier, 1814.

91. Aldrich J. R. A. Fisher on Bayes and Bayes' Theorem. *Bayesian Analysis*. 2008;3(1):161–70. doi:10.1214/08-BA306

92. Zampieri FG, Casey JD, Shankar-Hari M, et al. Using bayesian methods to augment the interpretation of critical care trials. an overview of theory and example reanalysis of the alveolar recruitment for acute respiratory distress syndrome trial. *American Journal of Respiratory and Critical Care Medicine*. 2021;203(5):543–52. doi:10.1164/rccm.202006-2381CP

93. van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. *Nature Reviews: Methods Primers*. 2021;1(1):1–26. doi:10.1038/s43586-020-00001-2

94. van Zwet E. A default prior for regression coefficients. *Statistical Methods in Medical Research*. 2019;28(12):3799–807. doi:10.1177/0962280218817792

95. Gelman A, Yao Y. Holes in Bayesian Statistics. *arXiv*. 2020. doi: 10.1088/1361-6471/abc3a5 doi:10.1088/1361-6471/abc3a5

96. Klitgaard TL, Schjørring OL, Lange T, et al. Bayesian and heterogeneity of treatment effect analyses of the HOT-ICU trial – a secondary analysis protocol. *Acta Anaesthesiologica Scandinavica*. 2020;9(Oct):1376–81. doi:10.1111/aas.13669

97. May TL, Ruthazer R, Riker RR, et al. Early withdrawal of life support after resuscitation from cardiac arrest is common and may result in additional deaths. *Resuscitation*. 2019;139(June):1–6. doi:10.1016/j.resuscitation.2019.02.031

98. Sung L, Hayden J, Greenberg ML, et al. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *Journal of Clinical Epidemiology*. 2005;58(3):261–8. doi:10.1016/j.jclinepi.2004.08.010

99. Combes A, Hajage D, Capellier G, et al. Extracorporeal Membrane Oxygenation for Severe Acute Respiratory Distress Syndrome. *New England Journal of Medicine*. 2018;378(21):1965–75. doi:10.1056/nejmoa1800385

100. Goligher EC, Tomlinson G, Hajage D, et al. Extracorporeal Membrane Oxygenation for Severe Acute Respiratory Distress Syndrome and Posterior Probability of Mortality Benefit in a Post Hoc Bayesian Analysis of a Randomized Clinical Trial. *JAMA - Journal of the American Medical Association*. 2018;320(21):2251–9. doi:10.1001/jama.2018.14276

101. Hernández G, Ospina-Tascón GA, Damiani LP, et al. Effect of a Resuscitation Strategy Targeting Peripheral Perfusion Status vs Serum Lactate Levels on 28-Day Mortality among Patients with Septic Shock: The ANDROMEDA-SHOCK Randomized Clinical Trial. *JAMA - Journal of the American Medical Association*. 2019;321(7):654–64. doi:10.1001/jama.2019.0071

102. Zampieri FG, Costa EL, Iwashyna TJ, et al. Heterogeneous effects of alveolar recruitment in acute respiratory distress syndrome: a machine learning reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome Trial. *British Journal of Anaesthesia*. 2019;123(1):88–95. doi:10.1016/j.bja.2019.02.026

103. Cavalcanti AB, Suzumura ÉA, Laranjeira LN, et al. Effect of lung recruitment and titrated Positive End-Expiratory Pressure (PEEP) vs low PEEP on mortality in patients with acute respiratory distress syndrome - A randomized clinical trial. *JAMA - Journal of the American Medical Association*. 2017;318(14):1335–45. doi:10.1001/jama.2017.14171

104. Zampieri FG, Damiani LP, Bakker J, et al. Effects of a Resuscitation Strategy Targeting Peripheral Perfusion Status versus Serum Lactate Levels among

Patients with Septic Shock A Bayesian Reanalysis of the ANDROMEDA-SHOCK Trial. *American Journal of Respiratory and Critical Care Medicine*. 2020;201(4):423–9. doi:10.1164/rccm.201905-0968OC

105. Krag M, Marker S, Perner A, et al. Pantoprazole in Patients at Risk for Gastrointestinal Bleeding in the ICU. *New England Journal of Medicine*. 2018;379(23):2199–208. doi:10.1056/NEJMoa1714919

106. Granholm A, Marker S, Krag M, et al. Heterogeneity of treatment effect of prophylactic pantoprazole in adult ICU patients: a post hoc analysis of the SUP-ICU trial. *Intensive Care Medicine*. 2020;46:717–29. doi:10.1007/s00134-019-05903-8

107. Hamilton FW, Lee T, Arnold DT, et al. Is convalescent plasma futile in COVID-19? A Bayesian re-analysis of the RECOVERY randomized controlled trial. *International Journal of Infectious Diseases*. 2021;109:114–7. doi:10.1016/j.ijid.2021.06.034

108. Abani O, Abbas A, Abbas F, et al. Convalescent plasma in patients admitted to hospital with COVID-19 (RECOVERY): a randomised controlled, open-label, platform trial. *The Lancet*. 2021;397(10289):2049–59. doi:10.1016/S0140-6736(21)00897-7

109. Investigators TR-C. Interleukin-6 Receptor Antagonists in Critically Ill Patients with Covid-19. *New England Journal of Medicine*. 2021;February 2:1–12. doi:10.1056/nejmoa2100433

110. Angus DC, Berry S, Lewis RJ, et al. The remap-cap (Randomized embedded multifactorial adaptive platform for community-acquired pneumonia) Study rationale and design. *Annals of the American Thoracic Society*. 2020;17(7):879–91. doi:10.1513/AnnalsATS.202003-192SD

111. Investigators TWC for the R-C. Effect of Hydrocortisone on Mortality and Organ Support in Patients with Severe COVID-19: The REMAP-CAP COVID-19 Corticosteroid Domain Randomized Clinical Trial. *JAMA - Journal of the American Medical Association*. 2020;324(13):1317–29. doi:10.1001/jama.2020.17022

112. Kendall M J. Designing a research project: randomised controlled trials and their principles. *Emergency Medicine Journal*. 2003;20:164–8.

113. Altman DG. Randomisation: Essential for reducing bias. *British Medical Journal*. 1991;302(6791):1481–2.

114. Begg C, Cho M, Eastwood S, et al. Improving the Quality of Reporting of Randomized Controlled Trials - The CONSORT Statement. *JAMA - Journal of the American Medical Association*. 1996;276(8):637–9. doi:10.1001/jama.276.8.637

115. Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: updated guidelines for reporting parellel group randomised trials. *BMC Medicine*. 2010;8:18:1–9. doi:doi: 10.1186/1741-7015-8-18

116. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*. 2010;63(8):e1–37. doi:10.1016/j.jclinepi.2010.03.004

117. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases*. 1967;20(8):637–48. doi:10.1016/0021-9681(67)90041-0

118. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues in Clinical Neuroscience*. 2011;13:217–24. doi:10.31887/dcns.2011.13.2/npatsopoulos

119. Califf RM, Sugarman J. Exploring the ethical and regulatory issues in pragmatic clinical trials. *Clinical Trials*. 2015;12(5):436–41. doi:10.1177/1740774515598334

120. Christensen E. Methodology of superiority vs. equivalence trials and non-inferiority trials. *Journal of Hepatology*. 2007;46(5):947–54. doi:10.1016/j.jhep.2007.02.015

121. Garrett AD. Therapeutic equivalence: Fallacies and falsification. *Statistics in Medicine*. 2003;22(5):741–62. doi:10.1002/sim.1360

122. Granholm A, Alhazzani W, Derde LPG, et al. Randomised clinical trials in critical care : past , present and future. *Intensive Care Medicine*. 2021. doi: 10.1007/s00134-021-06587-9 doi:10.1007/s00134-021-06587-9

123. Schjørring OL, Perner A, Wetterslev J, et al. Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU)—Protocol for a randomised clinical trial comparing a lower vs a higher oxygenation target in adults with acute hypoxaemic respiratory failure. *Acta Anaesthesiologica Scandinavica*. 2019;63(7):956–65. doi:10.1111/aas.13356

124. Schjørring OL, Klitgaard TL, Perner A, et al. The Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial: Detailed statistical analysis plan.

Acta Anaesthesiologica Scandinavica. 2020;64(6):847–56. doi:10.1111/aas.13569

125. Altman DG, Bland JM. Absence of evidence is not evidence of abscence. *British Medical Journal*. 1995;311(August 19):485.

126. Higgins J, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 [Internet]. 2021 Available from: www.training.cochrane.org/handbook

127. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*. 2011;343(d5928):1–9. doi:10.1136/bmj.d5928

128. Sterne JAC, Savović J, Page MJ, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *British Medical Journal*. 2019;366(l4898):1–8. doi:10.1136/bmj.l4898

129. Schünemann H, Brożek J, Guyatt GH, et al. GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. [Internet]. Available from: https://gdt.gradepro.org/app/handbook/handbook.html

130. Garattini S, Jakobsen JC, Wetterslev J, 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. doi:10.1016/j.ejim.2016.03.020

131. Klitgaard TL, Schjørring OL, Nielsen FM, et al. Letter to the editor: Serious methodological concerns about a recently published meta-analysis on oxygen therapy. *Journal of Intensive Care*. 2021;9(72):1–5. doi:10.1186/s40560-021-00573-5

132. Schjørring OL, Klitgaard TL, Perner A, et al. Lower or Higher Oxygenation Targets for Acute Hypoxemic Respiratory Failure. *New England Journal of Medicine*. 2021;384:1301–11. doi:10.1056/NEJMoa2032510

133. Klitgaard TL, Schjørring OL, Lange T, et al. Lower versus higher oxygenation targets in critically ill patients with severe hypoxaemia: secondary Bayesian analysis to explore heterogeneous treatment effects in the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial. *British Journal of Anaesthesia*. 2021;128(1):55–64. doi:10.1016/j.bja.2021.09.010

134. Klitgaard TL, Schjørring OL, Nielsen FM, et al. Higher versus lower fractions

of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (Updated review) - draft. *Cochrane Database of Systematic Reviews*. 2022;:CD012631. doi:10.1002/14651858.CD012631.pub3

135. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research*. 2011;20(10):1727–36. doi:10.1007/s11136-011-9903-x

136. Randolph C, Tierney MC, Mohr E, et al. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Preliminary Clinical Validity. *Journal of Clinical and Experimental Neuropsychology*. 1998;20(3):310–9.

137. Crescioli E, Riis J, Weinreich UM, et al. Long-term cognitive and pulmonary functions following a lower versus a higher oxygenation target in the HOT-ICU trial: protocol and statistical analysis plan. *Acta Anaesthesiologica Scandinavica*. 2021;(October):1–6. doi:10.1111/aas.13995

138. Fergusson D. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. *British Medical Journal*. 2002;325:652–4. doi:10.1136/bmj.325.7365.652

139. Krag M, Perner A, Wetterslev J, et al. Prevalence and outcome of gastrointestinal bleeding and use of acid suppressants in acutely ill adult intensive care patients. *Intensive Care Medicine*. 2015;41(5):833–45. doi:10.1007/s00134-015-3725-1

140. Jakobsen JC, Wetterslev J, Winkel P, et al. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Medical Research Methodology*. 2014

141. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. *Annals of Internal Medicine*. 1999;130(12):1005–13. doi:10.7326/0003-4819-130-12-199906150-00019

142. Jakobsen JC, Wetterslev J, Winkel P, et al. Bayes factor calculator [Internet]. Available from: www.ctu.dk/tools-and-links/bayes-factor-calculation.aspx

143. Vincent J-L, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine*. 1996;22(7):707–10. doi:10.1007/BF01709751

144. International conference on harmonisation; guidance on statistical principles

for clinical trials; availability -FDA Notice. *Federal register*. 1998;Sep 16(63(179)):49583–98.

145. Carpenter B, Gelman A, Hoffman MD, et al. Stan: A probabilistic programming language. *Journal of Statistical Software*. 2017; 76. doi: 10.18637/jss.v076.i01 doi:10.18637/jss.v076.i01

146. Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*. 2017; 80. doi: 10.18637/jss.v080.i01 doi:10.18637/jss.v080.i01

147. Bürkner PC. Advanced Bayesian multilevel modeling with the R package brms. *R Journal*. 2018;10(1):395–411. doi:10.32614/rj-2018-017

148. Team SD. RStan: the R interface to Stan. 2020 Available from: http://mc-stan.org/

149. Barbateskovic M, Schjørring OLL, Jakobsen JC, et al. Higher versus lower inspiratory oxygen fraction or targets of arterial oxygenation for adult intensive care patients (Protocol). *Cochrane Database of Systematic Reviews*. 2017 Apr 27. doi: 10.1002/14651858.CD012631 doi:10.1002/14651858.CD012631

150. 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

151. Wetterslev J, Jakobsen JC, Gluud C. Trial Sequential Analysis in systematic reviews with meta-analysis. *BMC Medical Research Methodology*. 2017;17(1):1–18. doi:10.1186/s12874-017-0315-7

152. Thorlund K, Engstrøm J, Wetterslev J, et al. User Manual for Trial Sequential Analysis (TSA) [Internet]. 2011. Available from: https://ctu.dk/tsa/learn-more/

153. Wetterslev J, Thorlund K, Brok J, et al. Estimating required information size by quantifying diversity in random-effects model meta-analyses. *BMC Medical Research Methodology*. 2009;9:1–12. doi:10.1186/1471-2288-9-86

154. Wetterslev J, Thorlund K, Brok J, et al. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. *Journal of Clinical Epidemiology*. 2008;61(1):64–75. doi:10.1016/j.jclinepi.2007.03.013

155. Imberger G, Gluud C, Boylan J, et al. 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. doi:10.1213/ANE.0000000000000892

156. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *British Medical Journal*. 2003;327(7414):557–60.

157. GRADEpro GDT [Internet]. 2015 Available from: gradepro.org

158. Rasmussen BS, Klitgaard TL, Perner A, et al. Oxygenation targets in ICU patients with COVID-19: a post-hoc sub-group analysis of the HOT-ICU trial. *Acta Anaesthesiologica Scandinavica*. 2021;(June):1–9. doi:10.1111/aas.13977

159. Barbateskovic M, Schjørring OL, Krauss SR, et al. Higher vs Lower Oxygenation Strategies in Acutely Ill Adults. *Chest*. 2020;159(1):154–73. doi:10.1016/j.chest.2020.07.015

160. Jun J, Sun L, Wang Y, et al. Invasive Mechanical Ventilation With High Concentration Oxygen Therapy for Aecopd Patients With Acute Myocardial Infarction. *Chest*. 2019;156(4):A958. doi:10.1016/j.chest.2019.08.886

161. Martin DS, McNeil M, Brew-Graves C, et al. A feasibility randomised controlled trial of targeted oxygen therapy in mechanically ventilated critically ill patients. *Journal of the Intensive Care Society*. 2021. doi: 10.1177/17511437211010031 doi:10.1177/17511437211010031

162. Yang X, Shang Y, Yuan S. Low versus high pulse oxygen saturation directed oxygen therapy in critically ill patients: A randomized controlled pilot study. *Journal of Thoracic Disease*. 2019;11(10):4234–40. doi:10.21037/jtd.2019.09.66

163. Panwar R. The unknowns about oxygen therapy in critically ill patients. *Journal of Thoracic Disease*. 2016;8(11):E1543–6. doi:10.21037/jtd.2016.11.85

164. International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH) adopts co. *International Digest of Health Legislation*. 1997;48(2):231–4.

165. JB Cabello, Burls A, Emparanza J, et al. Oxygen therapy for acute myocardial infarction (Review). *Cochrane Database of Systematic Reviews*. 2016;1–

71(12):Art. No.: CD007160. doi:10.1002/14651858.CD007160.pub4

166. Sepehrvand N, James SK, Stub D, et al. Effects of supplemental oxygen therapy in patients with suspected acute myocardial infarction: A meta-analysis of randomised clinical trials. *Heart*. 2018;104(20):1691–8. doi:10.1136/heartjnl-2018-313089

167. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence - Inconsistency. *Journal of Clinical Epidemiology*. 2011;64(12):1294–302. doi:10.1016/j.jclinepi.2011.03.017

168. Hultcrantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *Journal of Clinical Epidemiology*. 2017;87(1):4–13. doi:10.1016/j.jclinepi.2017.05.006

169. MEGA-ROX trial (ANZICS reg.no. CTG1920-01) [Internet]. [cited 2021 Apr 14]. Available from: https://www.anzics.com.au/current-active-endorsed-research/mega-rox/

170. UK-ROX (ICNARC project number: NIHR130508) [Internet]. [cited 2021 Apr 14]. Available from: https://www.icnarc.org/Our-Research/Studies/Uk-Rox

171. Kratz A, Ferraro M, Sluss PM, et al. Case Records of the Massachusetts General Hospital. Laboratory Reference Values. *New England Journal of Medicine*. 2004;354:145–1548. doi:10.1056/NEJMcpc049016

172. Pastene B, Leone M. ICU Management & Practice - Oxygen Therapy. *Journal of Patient Safety*. 2020;21(3):130–2.

173. Yang W, Zhang L. Observation of the curative effect of conservative oxygen therapy in mechanical ventilation of patients with severe pneumonia. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue*. 2021;33(9):1069–73. doi:10.3760/cma.j.cn121430-20210617-00902

174. Brower R, Lanken P, MacIntyre N, 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.

175. Singer M, Deutschman CS, Seymour C, et al. The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA - Journal of the American Medical Association*. 2016;315(8):801–10. doi:10.1001/jama.2016.0287

176. Hochberg CH, Semler MW, Brower RG. Oxygen toxicity in critically ill adults.

American Journal of Respiratory and Critical Care Medicine. 2021;204(6):632–41. doi:10.1164/rccm.202102-0417CI

177. Rhodes KM, Turner RM, Savović J, et al. Between-trial heterogeneity in meta-analyses may be partially explained by reported design characteristics. *Journal of Clinical Epidemiology*. 2018;95:45–54. doi:10.1016/j.jclinepi.2017.11.025

178. Iwashyna TJ, Burke JF, Sussman JB, et al. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *American Journal of Respiratory and Critical Care Medicine*. 2015;192(9):1045–51. doi:10.1164/rccm.201411-2125CP

179. Kent DM, Paulus JK, Van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) statement. *Annals of Internal Medicine*. 2020;172(1):35–45. doi:10.7326/M18-3667

180. Young P, Mackle D, Bellomo R, et al. Conservative oxygen therapy for mechanically ventilated adults with sepsis: a post hoc analysis of data from the intensive care unit randomized trial comparing two approaches to oxygen therapy (ICU-ROX). *Intensive care medicine*. 2020;46(1):17–26. doi:10.1007/s00134-019-05857-x

181. Demiselle J, Wepler M, Hartmann C, et al. Hyperoxia toxicity in septic shock patients according to the Sepsis-3 criteria: a post hoc analysis of the HYPER2S trial. *Annals of Intensive Care*. 2018;8(1):1–10. doi:10.1186/s13613-018-0435-1

182. Evans L, Rhodes A, Alhazzani W, et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock 2021. 2021. doi:10.1097/CCM.0000000000005337

183. Perner A, De Jong A, Shankar-Hari M. Trials on oxygen supplementation in sepsis: better late than never. *Intensive Care Medicine*. 2020;46(1):116–8. doi:10.1007/s00134-019-05874-w

184. Young P, Mackle D, Bellomo R, et al. Conservative oxygen therapy for mechanically ventilated adults with suspected hypoxic ischaemic encephalopathy. *Intensive Care Medicine*. 2020;46(12):2411–22. doi:10.1007/s00134-020-06196-y

185. Wang CH, Chang WT, Huang CH, et al. The effect of hyperoxia on survival following adult cardiac arrest: A systematic review and meta-analysis of observational studies. *Resuscitation*. 2014;85(9):1142–8.

doi:10.1016/j.resuscitation.2014.05.021

186. Young PJ, Bailey M, Bellomo R, et al. Conservative or liberal oxygen therapy in adults after cardiac arrest: An individual-level patient data meta-analysis of randomised controlled trials. *Resuscitation*. 2020;157(September):15–22. doi:10.1016/j.resuscitation.2020.09.036

187. Austin MA, Wills KE, Blizzard L, et al. Effect of high flow oxygen on mortality in chronic obstructive pulmonary disease patients in prehospital setting: Randomised controlled trial. *British Medical Journal*. 2010;341:c5462. doi:10.1136/bmj.c5462

188. Echevarria C, Steer J, Wason J, et al. Oxygen therapy and inpatient mortality in COPD exacerbation. *Emergency Medicine Journal*. 2021;38(3):170–7. doi:10.1136/emermed-2019-209257

189. Kopsaŕtis Z, Carson-Chahhoud K V., Austin MA, et al. Oxygen therapy in the pre-hospital setting for acute exacerbations of chronic obstructive pulmonary disease (Review). *Cochrane Database of Systematic Reviews*. 2020;(1):Art. No.: CD005534. doi:10.1016/j.ccm.2020.05.002

190. Cumpstey A, Oldman AH, Martin D, et al. Oxygen targets in the intensive care unit during mechanical ventilation for acute respiratory distress syndrome: a rapid review. *Cochrane database of systematic reviews (Online)*. 2020. doi: 10.1002/14651858.CD013708 doi:10.1002/14651858.CD013708

191. Saha R, Assouline B, Mason G, et al. Impact of differences in acute respiratory distress syndrome randomised controlled trial inclusion and exclusion criteria: systematic review and meta-analysis. *British Journal of Anaesthesia*. 2021;127(1):85–101. doi:10.1016/j.bja.2021.02.027

192. Stub D, Smith K, Bernard S, et al. Air versus oxygen in ST-segment-elevation myocardial infarction. *Circulation*. 2015;131(24):2143–50. doi:10.1161/CIRCULATIONAHA.114.014494

193. Hofmann R, James SK, Jernberg T, et al. Oxygen Therapy in Suspected Acute Myocardial Infarction. *New England Journal of Medicine*. 2017;377(13):1240–9. doi:10.1056/NEJMoa1706222

194. Mokhtari A, Akbarzadeh M, Sparv D, et al. Oxygen therapy in patients with ST elevation myocardial infarction based on the culprit vessel: Results from the randomized controlled SOCCER trial. *BMC Emergency Medicine*. 2020;20(1):1–10. doi:10.1186/s12873-020-00309-y

195. Zhang R, Zhu Y, Zhang M, et al. Oxygen therapy versus conservative therapy in suspected uncomplicated myocardial infarction without hypoxemia: A meta-analysis of randomized controlled studies. *Hong Kong Journal of Emergency Medicine*. 2021;28(6):367–78. doi:10.1177/1024907919894416

196. Roffe C, Nevatte T, Sim J, 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 - Journal of the American Medical Association*. 2017;318(12):1125–35. doi:10.1001/jama.2017.11463

197. Baekgaard JS, Isbye D, Ottosen CI, et al. Restrictive vs liberal oxygen for trauma patients-the TRAUMOX1 pilot randomised clinical trial. *Acta Anaesthesiologica Scandinavica*. 2019;(February):1–9. doi:10.1111/aas.13362

198. Christensen MA, Steinmetz J, Velmahos G, et al. Supplemental oxygen therapy in trauma patients: An exploratory registry-based study. *Acta Anaesthesiologica Scandinavica*. 2021;65(7):967–78. doi:10.1111/aas.13829

199. Kondo Y, Gibo K, Abe T, et al. Association of prehospital oxygen administration and mortality in severe trauma patients (PROMIS): A nationwide cohort study. *Medicine*. 2019;98(27):e16307. doi:10.1097/MD.0000000000016307

200. Douin DJ, Anderson EL, Dylla L, et al. Association Between Hyperoxia, Supplemental Oxygen, and Mortality in Critically Injured Patients. *Critical Care Explorations*. 2021;3(5):e0418. doi:10.1097/cce.0000000000000418

201. Ebmeier SJ, Barker M, Bacon M, et al. A Two Centre Observational Study of Simultaneous Pulse Oximetry and Arterial Oxygen Saturation Recordings in Intensive Care Unit Patients. *Anaesthesia and Intensive Care*. 2018;46(3):297–303. doi:10.1177/0310057X1804600307

202. Schjørring OL, Rasmussen BS. The paramount parameter: arterial oxygen tension versus arterial oxygen saturation as target in trials on oxygenation in intensive care. *Critical Care*. 2018;22(1):9–11. doi:10.1186/s13054-018-2257-9

203. Goodford PJ, Norrington FE, Paterson RA, et al. The effect of 2,3-diphosphoglycerate in the oxygen dissociation curve of human haemoglobin. *Journal of Physiology*. 1977;273:631–45.

204. Siggaard-Andersen O, Wimberley PD, Göthgen I, et al. A mathematical model of the hemoglobin-oxygen dissociation curve of human blood and of the

oxygen partial pressure as a function of temperature. *Clinical Chemistry*. 1984;30(10):1646–51. doi:10.1093/clinchem/30.10.1646

205. Anstey C. A new model for the oxyhaemoglobin dissociation curve. *Anaesthesia and Intensive Care*. 2003;31(4):349–87.

206. Patel S, Jose A, Mohiuddin SS. Physiology, Oxygen Transport And Carbon Dioxide Dissociation Curve [Internet]. 2021 Available from: https://www.ncbi.nlm.nih.gov/books/NBK539815/

207. Kaufman DP, Kandle PF, Murray I, et al. Physiology, Oxyhemoglobin Dissociation Curve [Internet]. 2021 Available from: https://www.ncbi.nlm.nih.gov/books/NBK499818/

208. Thabane L, Mbuagbaw L, Zhang S, et al. A tutorial on sensitivity analyses in clinical trials: The what, why, when and how. *BMC Medical Research Methodology*. 2013;13(1):1. doi:10.1186/1471-2288-13-92

209. Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. *New England Journal of Medicine*. 2017;377(14):1391–8. doi:10.1056/NEJMsm1605385

210. Hróbjartsson A, Emanuelsson F, Thomsen ASS, et al. 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. doi:10.1093/ije/dyu115

211. Savović J, Turner RM, Mawdsley D, 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. doi:10.1093/aje/kwx344

212. Sjoding MW, Dickson RP, Iwashyna TJ, et al. Racial Bias in Pulse Oximetry Measurement Michael. *New England Journal of Medicine*. 2020;383(25):2477–8. doi:10.1056/NEJMc2029240

213. Bellani G, Laffey JG, Pham T, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA - Journal of the American Medical Association*. 2016;315(8):788–800. doi:10.1001/jama.2016.0291

214. Perner A, Haase N, Guttormsen AB, et al. Hydroxyethyl starch 130/0.42 versus Ringer's acetate in severe sepsis. *New England Journal of Medicine*. 2012;367(2):124–34. doi:10.1177/175114371301400116

215. Holst LB, Haase N, Wetterslev J, et al. Lower versus higher haemoglobin threshold for transfusion in septic shock. *New England Journal of Medicine*. 2014;371(15):1381–91. doi:10.1177/1751143715607722

216. Munch MW, Myatra SN, Vijayaraghavan BKT, et al. Effect of 12 mg vs 6 mg of Dexamethasone on the Number of Days Alive without Life Support in Adults with COVID-19 and Severe Hypoxemia: The COVID STEROID 2 Randomized Trial. *JAMA - Journal of the American Medical Association*. 2021;326(18):1807–17. doi:10.1001/jama.2021.18295

217. Agents Intervening Against Delirium in Intensive Care Unit (AID-ICU). *ClinicalTrials.gov Identifier: NCT03392376*.

218. Handling Oxygenation Targets in COVID-19 (HOT-COVID). *ClinicalTrials.gov Identifier: NCT04425031*.

219. Goal directed fluid removal with furosemide in intensive care patients with fluid overload – A randomised, blinded, placebo-controlled trial (GODIF). *ClinicalTrials.gov Identifier: NCT04180397*.

220. The Conservative vs. Liberal Approach to Fluid Therapy of Septic Shock in Intensive Care Trial (CLASSIC). *ClinicalTrials.gov Identifier: NCT03668236*.

221. Fuchs PA, Czech IJ, Krzych ŁJ. Mortality prediction using SOFA score in critically ill surgical and non-surgical patients: Which parameter is the most valuable? *Medicina*. 2020;56(6):1–8. doi:10.3390/medicina56060273

222. Schoe A, Bakhshi-Raiez F, De Keizer N, et al. Mortality prediction by SOFA score in ICU-patients after cardiac surgery; Comparison with traditional prognostic-models. *BMC Anesthesiology*. 2020;20(1):1–8. doi:10.1186/s12871-020-00975-2

223. Granholm A, Møller MH, Krag M, et al. Predictive performance of the simplified acute physiology score (SAPS) II and the initial sequential organ failure assessment (SOFA) score in acutely ill intensive care patients: Post-hoc analyses of the SUP-ICU inception cohort study. *PLoS ONE*. 2016; 11. doi: 10.1371/journal.pone.0168948 doi:10.1371/journal.pone.0168948

224. Ferreira FL, Bota DP, Mélot C, et al. Serial Evaluation of the SOFA Score. *JAMA - Journal of the American Medical Association*.;286(14):1754–8.

225. 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 Medical*

Research Methodology. 2014;14(1):1–12. doi:10.1186/1471-2288-14-25

226. Barbateskovic M, Koster TM, Eck RJ, et al. A new tool to assess Clinical Diversity In Meta-analyses (CDIM) of interventions. *Journal of Clinical Epidemiology*. 2021;135:29–41. doi:10.1016/j.jclinepi.2021.01.023

APPENDICES

Appendix A. Paper I

Lower or Higher Oxygenation Targets for Acute Hypoxemic Respiratory Failure
Schjørring OL*, **Klitgaard TL***, Perner A, et al. for the HOT-ICU trial group.
New England Journal of Medicine. 2021. 384(14):1301-1311
(Online ahead of print January 20, 2021)
*Shared first authorship
doi: 10.1056/NEJMoa2032510

Link: nejm.org/doi/full/10.1056/NEJMoa2032510

ORIGINAL ARTICLE

Lower or Higher Oxygenation Targets for Acute Hypoxemic Respiratory Failure

O.L. Schjørring, T.L. Klitgaard, A. Perner, J. Wetterslev, T. Lange, M. Siegemund, M. Bäcklund, F. Keus, J.H. Laake, M. Morgan, K.M. Thormar, S.A. Rosborg, J. Bisgaard, A.E.S. Erntgaard, A.-S.H. Lynnerup, R.L. Pedersen, E. Crescioli, T.C. Gielstrup, M.T. Behzadi, L.M. Poulsen, S. Estrup, J.P. Laigaard, C. Andersen, C.B. Mortensen, B.A. Brand, J. White, I.-L. Jarnvig, M.H. Møller, L. Quist, M.H. Bestle, M. Schønemann-Lund, M.K. Kamper, M. Hindborg, A. Hollinger, C.E. Gebhard, N. Zellweger, C.S. Meyhoff, M. Hjort, L.K. Bech, T. Grøfte, H. Bundgaard, L.H.M. Østergaard, M.A. Thyø, T. Hildebrandt, B. Uslu, C.G. Sølling, N. Møller-Nielsen, A.C. Brøchner, M. Borup, M. Okkonen, W. Dieperink, U.G. Pedersen, A.S. Andreasen, L. Buus, T.N. Aslam, R.R. Winding, J.C. Schefold, S.B. Thorup, S.A. Iversen, J. Engstrøm, M.-B.N. Kjær, and B.S. Rasmussen, for the HOT-ICU Investigators*

ABSTRACT

BACKGROUND

Patients with acute hypoxemic respiratory failure in the intensive care unit (ICU) are treated with supplemental oxygen, but the benefits and harms of different oxygenation targets are unclear. We hypothesized that using a lower target for partial pressure of arterial oxygen (Pao_2) would result in lower mortality than using a higher target.

METHODS

In this multicenter trial, we randomly assigned 2928 adult patients who had recently been admitted to the ICU (≤12 hours before randomization) and who were receiving at least 10 liters of oxygen per minute in an open system or had a fraction of inspired oxygen of at least 0.50 in a closed system to receive oxygen therapy targeting a Pao_2 of either 60 mm Hg (lower-oxygenation group) or 90 mm Hg (higher-oxygenation group) for a maximum of 90 days. The primary outcome was death within 90 days.

RESULTS

At 90 days, 618 of 1441 patients (42.9%) in the lower-oxygenation group and 613 of 1447 patients (42.4%) in the higher-oxygenation group had died (adjusted risk ratio, 1.02; 95% confidence interval, 0.94 to 1.11; P=0.64). At 90 days, there was no significant between-group difference in the percentage of days that patients were alive without life support or in the percentage of days they were alive after hospital discharge. The percentages of patients who had new episodes of shock, myocardial ischemia, ischemic stroke, or intestinal ischemia were similar in the two groups (P=0.24).

CONCLUSIONS

Among adult patients with acute hypoxemic respiratory failure in the ICU, a lower oxygenation target did not result in lower mortality than a higher target at 90 days. (Funded by the Innovation Fund Denmark and others; HOT-ICU ClinicalTrials.gov number, NCT03174002.)

The authors' full names, academic degrees, and affiliations are listed in the Appendix. Address reprint requests to Dr. Rasmussen at the Department of Anesthesia and Intensive Care, Aalborg University Hospital, Hobrovej 18-21, DK-9000 Aalborg, Denmark, or at bodil.steen.rasmussen@rn.dk.

*A complete list of investigators in the HOT-ICU trial is provided in the Supplementary Appendix, available at NEJM.org.

Drs. Schjørring and Klitgaard contributed equally to this article.

This article was published on January 20, 2021, at NEJM.org.

DOI: 10.1056/NEJMoa2032510

Patients who are admitted to the intensive care unit (ICU) with acute hypoxemic respiratory failure often receive supplemental oxygen with a high fraction of inspired oxygen (F_{IO_2}), which results in a high partial pressure of arterial oxygen (P_{AO_2}). In some clinical trials, such therapy has been associated with increased mortality.[1-3] However, clinical practice guidelines give no recommendation for oxygenation targets in adult patients in the ICU owing to sparse evidence.[4-7]

In a small, multicenter, randomized trial involving patients undergoing mechanical ventilation in the ICU,[8] investigators found that targeting a peripheral oxygen saturation of 88 to 92%, as compared with a value of 96% or above, was feasible without evident harm. In a single-center, randomized trial,[9] patients in the ICU who were treated with a P_{AO_2} target of 70 to 100 mm Hg had lower mortality than those who were treated with a P_{AO_2} target of up to 150 mm Hg. In addition, a P_{AO_2} target of 55 to 80 mm Hg is often referred to as the standard of care in patients with acute respiratory distress syndrome (ARDS), as it was described in several trials performed by the ARDS Network.[10-12] The preference among clinicians for a lower oxygenation target in the ICU has been confirmed in a multinational survey, in which 80% of the respondents would accept a P_{AO_2} target of 60 mm Hg or lower in clinical trials.[13]

Recently, a systematic review and meta-analysis showed that lower oxygenation targets were preferable in acutely ill adults.[14] However, the Liberal Oxygenation versus Conservative Oxygenation in ARDS ($LOCO_2$) trial was stopped prematurely because of a higher frequency of mesenteric ischemia and a higher 90-day mortality in the lower-oxygenation group than in the higher-oxygenation group.[15] In the large Intensive Care Unit Randomized Trial Comparing Two Approaches to Oxygen Therapy (ICU-ROX), investigators found no between-group differences in the number of ventilator-free days or in mortality within 28 days.[16]

We conducted the Handling Oxygenation Targets in the ICU (HOT-ICU) trial to test the hypothesis that targeting a P_{AO_2} of 60 mm Hg would reduce 90-day mortality by 5 percentage points as compared with targeting a P_{AO_2} of 90 mm Hg in patients who were admitted to the ICU with hypoxemic respiratory failure.

Methods

Trial Design and Oversight

HOT-ICU was an investigator-initiated, multicenter, stratified, parallel-group clinical trial with centralized randomization and a computer-generated concealed assignment sequence, with permuted blocks of varying sizes, stratified according to trial site and the presence or absence of chronic obstructive pulmonary disease (COPD) or active hematologic cancer. From June 20, 2017, to August 3, 2020, patients were enrolled at 35 ICUs in Denmark, Switzerland, Finland, the Netherlands, Norway, the United Kingdom, and Iceland. Written informed consent for incapacitated patients without an available surrogate was temporarily obtained (from a doctor independent of the trial) until the patient regained capacity or a surrogate became available. If consent was withdrawn, we asked the patient or surrogate for permission to continue registration of trial data and to include the data in our analyses, in accordance with national regulations. Because of the nature of the trial, clinicians and patients or their surrogates were aware of the trial-group assignments.

The trial was designed and overseen by the steering committee. An independent data and safety monitoring committee, whose members were unaware of trial-group assignments, oversaw the trial and reviewed the planned interim analysis after 1464 patients had completed the 90-day follow-up. Trial data were reviewed at the sites by external monitors, in accordance with the Good Clinical Practice directive of the European Union, and centrally by staff from the coordinating center.

The trial protocol and the statistical analysis plan were published before the enrollment of the last patient in the trial[17,18] and are available in a single document with the full text of this article at NEJM.org. The protocol was approved by the relevant ethics committees, according to national regulations. The members of the steering committee wrote the first draft of the manuscript. All the authors vouch for the adherence of the trial to the protocol, for the accuracy and completeness of the data, and for the reporting of serious adverse events.

Patients

We screened adult patients (≥18 years of age) who were admitted to the ICU with hypoxemic respi-

ratory failure and who were receiving at least 10 liters of oxygen per minute in an open system or who had an F_{IO_2} of at least 0.50 in a closed system; all the patients had placement of an arterial line and were expected to receive supplementary oxygen therapy for at least 24 hours in the ICU. With these thresholds of oxygen supplementation, we assumed that the P_{aO_2}:F_{IO_2} ratio in all the patients would be below 300. We excluded patients who could not undergo randomization within 12 hours after ICU admission. All additional exclusion criteria are provided in the Supplementary Appendix, available at NEJM.org.

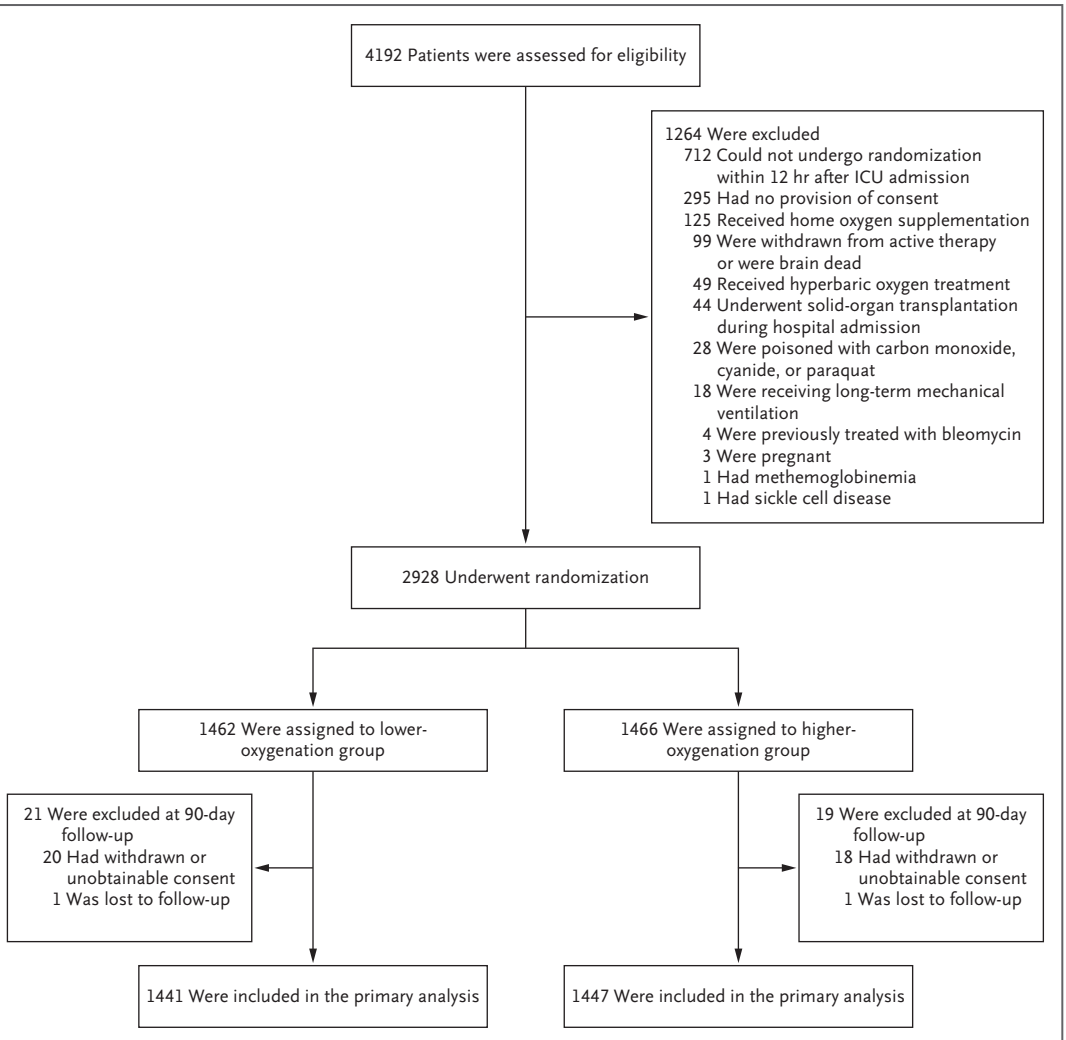


Figure 1. Screening, Randomization, and Follow-up.

Patients could have more than one reason for being excluded from the trial after screening. A total of 40 patients were excluded from the primary analysis after randomization because they or their surrogate did not allow the use of their data (17 in the lower-oxygenation group and 17 in the higher-oxygenation group) or the consent for the use of their data could not be obtained according to national regulations (3 patients and 1 patient, respectively); 1 patient in each group was lost to follow-up. Although 30 patients or surrogates (14 patients and 16 patients, respectively) did not want further data to be registered, mortality data were obtained from national registries, and these patients were included in the primary analysis; however, data regarding some secondary outcomes were missing. One patient in the lower-oxygenation group who had erroneously undergone randomization 5.5 hours after death was excluded from the primary analysis, and an additional patient underwent randomization. A supplemental analysis of the primary outcome that includes the erroneously randomized patient is provided in Table S9. ICU denotes intensive care unit.

Table 1. Characteristics of the Patients at Baseline.*

Characteristic	Lower-Oxygenation Group (N = 1453)	Higher-Oxygenation Group (N = 1457)
Median age (IQR) — yr	70 (60–77)	70 (60–77)
Male sex — no. (%)	925 (63.7)	946 (64.9)
Median interval between hospital admission and randomization (IQR) — days	1 (0–5)	1 (0–5)
Median interval between ICU admission and randomization (IQR) — hr	4 (2–7)	4 (2–7)
Coexisting illness — no. (%)		
Ischemic heart disease	205 (14.1)	205 (14.1)
Chronic heart failure	140 (9.6)	146 (10.0)
Active metastatic cancer	65 (4.5)	61 (4.2)
Long-term dialysis	19 (1.3)	28 (1.9)
Chronic obstructive pulmonary disease	277 (19.1)	286 (19.6)
Active hematologic cancer	82 (5.6)	86 (5.9)
Type of admission — no. (%)		
Medical	1248 (85.9)	1240 (85.1)
Elective surgery	18 (1.2)	21 (1.4)
Emergency surgery	187 (12.9)	196 (12.9)
Acute illness — no. (%)		
Pneumonia	838 (57.7)	836 (57.4)
Multiple trauma	24 (1.7)	29 (2.0)
Hemorrhagic or ischemic stroke	25 (1.7)	22 (1.5)
Traumatic brain injury	9 (0.6)	15 (1.0)
Myocardial infarction	84 (5.8)	99 (6.8)
Intestinal ischemia	27 (1.9)	41 (2.8)
Cardiac arrest	149 (10.3)	186 (12.8)
ARDS	178 (12.3)	195 (13.4)
Invasive ventilation		
Patients — no. (%)	834 (57.4)	870 (59.7)
Median tidal volume (IQR) — ml	499 (429–582)	499 (426–561)
Median end-expiratory pressure (IQR) — cm of water	9 (7–10)	10 (7–10)
Median peak pressure (IQR) — cm of water	25 (20–29)	25 (21–30)
Noninvasive ventilation or CPAP		
Patients — no. (%)	199 (13.7)	176 (12.1)
Median end-expiratory pressure (IQR) — cm of water	8 (6–9)	7 (5–8)
Open system — no. (%)	420 (28.9)	411 (28.2)
Median Pao_2 (IQR) — mm Hg	77.3 (65.3–93.8)	77.3 (62.3–93.0)
Median Sao_2 (IQR) — %†	94 (91–97)	95 (91–97)
Median Fio_2 (IQR) — fraction‡	0.70 (0.55–0.90)	0.70 (0.56–0.85)
Median Pao_2:Fio_2 ratio (IQR)		
In all systems	118.6 (88.8–157.5)	117.5 (90.0–153.8)
In closed systems	125.7 (91.6–165.0)	125.0 (94.7–163.5)

Table 1. (Continued.)

Characteristic	Lower-Oxygenation Group (N = 1453)	Higher-Oxygenation Group (N = 1457)
Median lactate level (IQR) — mmol/liter	1.8 (1.1–3.2)	1.7 (1.1–3.1)
Median lowest mean arterial pressure (IQR) — mm Hg§	59 (49–68)	58 (48–69)
Use of inotropes — no. (%)	33 (2.3)	37 (2.5)
Use of vasopressors		
Patients — no. (%)	800 (55.1)	791 (54.3)
Median highest dose of norepinephrine (IQR) — μg/kg/min	0.20 (0.10–0.40)	0.21 (0.10–0.40)
Median SOFA score (IQR)¶	9 (8–11)	9 (8–11)

* All baseline variables were missing for 9 patients in each group. ARDS denotes acute respiratory distress syndrome, CPAP continuous positive airway pressure, ICU intensive care unit, IQR interquartile range, and Pao_2 partial pressure of arterial oxygen.
† Values for arterial oxygen saturation (Sao_2) were not available for 191 patients because this measure was not included in blood gas analyses at one trial site.
‡ The fraction of inspired oxygen (Fio_2) in open systems was estimated with the use of standardized conversion tables.
§ Listed is lowest median value of the arterial pressure recorded during the 24 hours before randomization.
¶ Scores on the Sequential Organ Failure Assessment (SOFA) range from 0 to 24, with higher scores indicating more severe organ failure. Data were missing for 48 patients in the lower-oxygenation group and for 50 patients in the higher-oxygenation group.

INTERVENTION

Patients were randomly assigned in a 1:1 ratio to receive oxygen therapy targeting either a Pao_2 of 60 mm Hg (lower-oxygenation group) or a Pao_2 of 90 mm Hg (higher-oxygenation group) until a maximum of 90 days after randomization. The trial period included any readmissions to the ICU. We recorded the lowest and the highest Pao_2 in predefined 12-hour intervals, along with concomitant values of arterial oxygen saturation (Sao_2) and Fio_2. The oxygenation targets were achieved by adjustment of the Fio_2. In the two groups, deviations from the target of more than 7.5 mm Hg were accepted only in patients who had an Fio_2 of 0.21 or in those with an Fio_2 of 1.00. The oxygen-supplementation devices and ventilator settings were chosen by the clinicians. Ventilator settings were registered daily at 8 a.m. if either invasive or noninvasive ventilation or continuous positive airway pressure was being used. A schedule for the sampling of arterial blood gases was not mandated in the protocol, but we assumed that at least four measurements would be performed per day.[3] Since such measures of arterial blood gases were performed at varying times during the day, clinicians and nurses were instructed to monitor all patients with continuous measurement of peripheral oxygen saturation and to identify and maintain the saturation level at which the assigned Pao_2 was measured.

OUTCOME MEASURES

The primary outcome was death from any cause within 90 days after randomization. The secondary outcomes were the number of patients with one or more serious adverse events, which were defined as a new episode of shock, myocardial ischemia, cerebral ischemia, or intestinal ischemia; the percentage of days that patients were alive without life support, as defined by the absence of mechanical ventilation, renal-replacement therapy, or vasopressor or inotrope infusion; and the percentage of days that patients were alive after hospital discharge at the 90-day follow-up. (Additional details about the outcome measures are provided in the Supplementary Appendix.) Data regarding outcome measures were obtained from the patients' files by site investigators, who were aware of the trial-group assignments; data regarding 90-day mortality were also obtained from regional and national registries.

STATISTICAL ANALYSES

We estimated that the enrollment of 2928 patients would provide a power of 90% to detect a between-group difference of 5 percentage points in mortality at 90 days after randomization, which

would correspond to a 20% difference in relative risk at a two-sided alpha level of 5%. In making this determination, we assumed a 90-day mortality of 25% in the higher-oxygenation group on the basis of data from a study involving patients undergoing mechanical ventilation in five Danish ICUs.[3] Analyses of the primary and secondary outcomes were performed in the intention-to-treat population, which included all the patients who had undergone randomization, except those for whom consent was withdrawn or unobtainable.[19]

We compared dichotomous data between the two trial groups using a generalized linear model with a log-link and binomial error distribution with adjustment for stratification variables; results are reported as relative risks and risk differences with 95% confidence intervals for the primary outcome and with 98.75% confidence intervals for the secondary outcomes after adjustment for multiple comparisons.[18] Analysis of the primary outcome was supplemented with crude Kaplan–Meier plots and the calculation of a hazard ratio from a Cox proportional-hazards model with adjustment for stratification variables, as well as a Bayes factor calculation.[20] We used the Van Elteren test after adjustment only for the trial site to compare continuous data, since the assumptions of a Poisson distribution or a negative binomial distribution were not met.[21] Since the trial-group assignments could not be blinded, the analyses of the primary and secondary outcomes were performed with the oxygenation targets masked, and the steering committee wrote two abstracts assuming opposite group assignments before unblinding of the data (see the Supplementary Appendix). These two abstracts document the fully implemented blinding in the statistical analyses and in the main interpretation of the results. Statistical significance was indicated by a two-sided P value below 0.05 for the primary outcome and by a multiplicity-adjusted P value below 0.0125 for the three secondary outcomes.

We conducted a secondary analysis of the primary outcome in the intention-to-treat population using logistic regression (reported as odds ratios and 95% confidence intervals) after adjustment for the stratification variables and predefined risk factors at baseline: age, type of ICU admission, presence or absence of metastatic cancer, and the score on the Sequential Organ Failure Assessment (SOFA). (The SOFA score ranges from 0 to 24, as calculated from subscores ranging from 0 to 4 for each of six organ systems — respiration, coagulation, liver, cardiovascular, central nervous system, and renal — with higher scores indicating more severe organ failure.)[22]

We evaluated the primary outcome in subgroups that were defined according to the presence or absence of shock at the time of randomization, the use of invasive mechanical ventilation, COPD, traumatic brain injury, and cardiac arrest, along with the type of ICU admission (medical, elective surgery, or emergency surgery).[18] Details regarding the subgroup evaluations are provided in the Supplementary Appendix. A per-protocol analysis is also ongoing, so the results are not reported here. No imputations for missing data were performed, since the percentage of missing data was less than 5% for all outcomes.[23] All analyses were performed with the use of Stata statistical software, release 16 (StataNordic).

RESULTS

TRIAL POPULATION

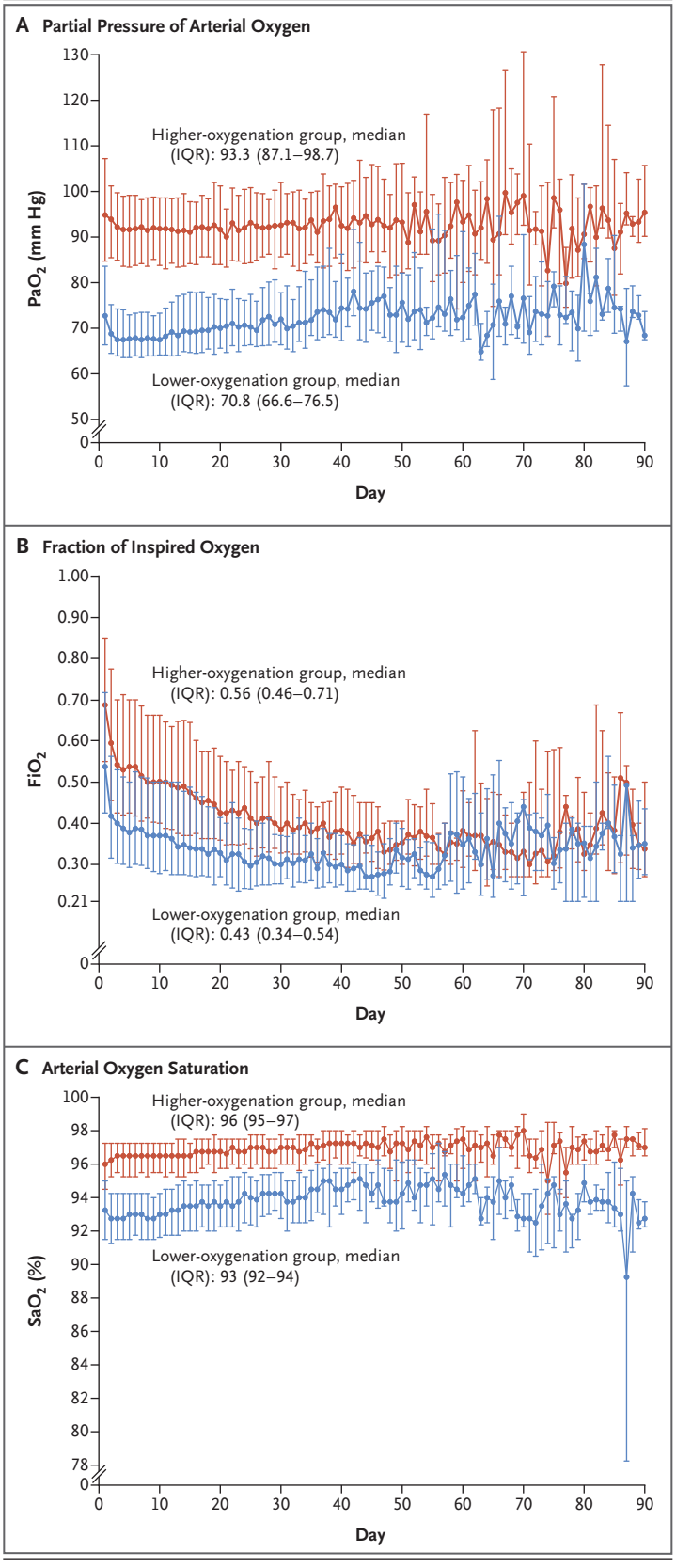
Of the 2928 patients who were enrolled in the trial, 1462 were assigned to the lower-oxygenation group and 1466 to the higher-oxygenation group. We obtained 90-day mortality data regarding 2888 patients (98.6%), which included 1441 patients in the lower-oxygenation group and 1447 patients in the higher-oxygenation group (Fig. 1). The trial groups had similar characteristics at baseline, except for the presence of cardiac arrest (Table 1).

OXYGENATION AND ICU INTERVENTIONS

During the 90-day intervention period, the recorded Pao_2 measurements were lower in the lower-oxygenation group than in the higher-oxygenation group, as were the corresponding Sao_2 and Fio_2 values (Fig. 2). The 12-hour highest and lowest Pao_2 measurements, with corresponding Sao_2 and Fio_2 values, are provided in Figures S1 through S3 in the Supplementary Appendix. The use of mechanical ventilation, prone positioning, inhaled vasodilators, extracorporeal membrane oxygenation, circulatory support, renal-replacement therapy, and blood transfusions were similar in the two groups. Data obtained daily at 8 a.m.

Figure 2. Values for Pao_2, Fio_2, and Sao_2, According to Oxygenation Strategy.

Shown are the median values of daily means of partial pressure of arterial oxygen (Pao_2) (Panel A), fraction of inspired oxygen (Fio_2) (Panel B), and arterial oxygen saturation (Sao_2) (Panel C) of the trial patients until a maximum of 90 days. The daily means were calculated from the 12-hour lowest and highest Pao_2 with concomitant values for Fio_2 and Sao_2. I bars represent interquartile ranges (IQR). Sao_2 values were not available in blood gas analyses from one site and were therefore missing for 191 patients. Data for patients according to day are provided in Table S1.



showed no substantial between-group differences regarding positive end-expiratory pressure, peak inspiratory pressure, or tidal volume among the patients who were undergoing invasive mechanical ventilation or in end-expiratory pressure among those who were undergoing noninvasive ventilation (Table S2).

OUTCOMES

At 90 days after randomization, 618 of 1441 patients (42.9%) in the lower-oxygenation group and 613 of 1447 patients (42.4%) in the higher-oxygenation group had died (risk ratio, 1.02; 95% confidence interval [CI], 0.94 to 1.11; P=0.64) (Table 2). Results were similar in the analysis after adjustment for baseline factors; the hazard ratio was similar as well after adjustment for stratification variables (Fig. 3). A Bayes factor that was substantially higher than 1 supported the finding of no effect of the intervention (see the Supplementary Appendix). The results of the subgroup analyses were similar to those in the primary analysis (Table S3).

At day 90, the percentage of days that patients were alive without life support and the percentage of days that patients were alive after hospital discharge did not differ significantly between the two groups (Table 2; absolute numbers and single components are provided in Tables S4, S5, and S6). Likewise, the number of patients with one or more serious adverse events did not differ significantly between the two groups (Table 2).

DISCUSSION

In this multicenter, randomized trial involving adult patients with acute hypoxemic respiratory failure in the ICU, we found that targeting a Pao_2 of 60 mm Hg rather than a Pao_2 of 90 mm Hg

Table 2. Primary and Secondary Outcomes.

Outcome	Lower-Oxygenation Group	Higher-Oxygenation Group	Risk Ratio (95% CI)*	Risk Difference (95% CI)*	Adjusted Odds Ratio (95% CI)	P Value
Primary outcome†						
Death by day 90 — no./total no. (%)	618/1441 (42.9)	613/1447 (42.4)				
Adjusted for stratification variables‡			1.02 (0.94 to 1.11)	0.63 (−2.92 to 4.17)		0.64
Adjusted for stratification and baseline variables§					1.06 (0.90 to 1.24)	0.50
Secondary outcomes¶						
Median percentage of days alive without life support (IQR)	87.8 (0.0–96.7)	84.4 (0.0–96.0)				0.10
Median percentage of days alive after hospital discharge (IQR)	55.6 (0.0–85.6)	50.0 (0.0–84.4)				0.67
Serious adverse events — no./total no. (%)	525/1453 (36.1)	555/1457 (38.1)	0.95 (0.84 to 1.07)	−1.6 (−6.0 to 2.8)		0.24
Shock	492/1453 (33.9)	521/1457 (35.8)				
Myocardial ischemia	14/1453 (1.0)	8/1457 (0.5)				
Ischemic stroke	19/1453 (1.3)	23/1457 (1.6)				
Intestinal ischemia	32/1453 (2.2)	29/1457 (2.0)				

* For serious adverse events, relative risk and risk difference are reported with 98.75% confidence intervals that have been adjusted for multiple comparisons. Risk differences are reported in percentage points.

† Data regarding the primary outcome were missing for 21 patients in the lower-oxygenation group and for 19 patients in the higher-oxygenation group.

‡ Stratification variables were the trial site and the presence or absence of chronic obstructive pulmonary disease or active hematologic cancer.

§ Baseline variables were age, presence or absence of active metastatic cancer, type of admission (medical, elective surgical, or emergency surgical), and the SOFA score, which ranges from 0 to 24, with higher scores indicating more severe organ failure.

¶ The percentage of days alive without life support was calculated as the number of days without the use of invasive ventilation, noninvasive ventilation, continuous positive airway pressure, vasopressor or inotropic infusion, or renal-replacement therapy, divided by the number of days alive within 90 days. The percentage of days alive after hospital discharge was calculated as the number of days alive and discharged from the hospital divided by the number of days alive within 90 days. Data were missing for 33 patients in each of the oxygenation groups. Absolute numbers and percentages are provided in Tables S6, S7, and S8.

did not result in better values for several key outcomes — including mortality, the percentage of days alive without life support, the percentage of days alive after hospital discharge, and serious adverse events — at 90 days. Our findings lend weight to the utility of conservative oxygen therapy in patients with acute hypoxemic respiratory failure, as compared with the results of the $LOCO_2$ trial.[15] At the same time, the results of our trial do not preclude the possibility of clinically important harm or benefit with a lower-oxygenation strategy in this population or in other types of critically ill patients. In the $LOCO_2$ trial, mesenteric ischemia occurred in five patients who were assigned to a Pao_2 target of 55 to 70 mm Hg and in no patients assigned to a Pao_2 target of 90 to 105 mm Hg. The overall incidence of intestinal ischemia in our trial (2.1%) was similar to that in the $LOCO_2$ trial (2.5%).[15] The $LOCO_2$ trial was stopped early after the inclusion of 201 patients with ARDS; at the time, there was no significant between-group difference in the primary outcome of mortality at day 28, but there was significantly higher 90-day mortality in the lower-oxygenation group. Although we recruited patients with acute hypoxemic respiratory failure regardless of the presence of ARDS, the baseline Pao_2:Fio_2 ratios were remarkably similar to those in the $LOCO_2$ trial.

Notably, we observed a 90-day mortality that was twice as high as had been hypothesized on the basis of data previously obtained in five Danish

ICUs.[3] The higher 90-day mortality in our trial may have been partially due to differences in the types of admissions. Acute medical conditions accounted for 85.5% of the admissions in our trial and for 37.3% of those in the cited cohort study, whereas emergency surgery accounted for 1.3% and 29.8%, respectively, and elective surgery for 13.2% and 32.6%, respectively. Furthermore, although only 12.8% of our patients were recorded as having ARDS at baseline, they had more severe hypoxemic respiratory failure than anticipated, with Pao_2:Fio_2 ratios in the range of those found in patients with moderate-to-severe ARDS. This degree of hypoxemia might also have contributed to the higher mortality observed in our trial. Accordingly, the present results may not be representative of outcomes in a lower-risk population.

In the ICU-ROX trial,[16] not all the patients had acute hypoxemic respiratory failure, as illustrated by a Pao_2:Fio_2 ratio at baseline that was twice as high as that both in our trial and in the $LOCO_2$ trial, as well as a lower Fio_2. The ICU-ROX trial showed no significant between-group differences in the number of ventilator-free days or in mortality at 90 days and 180 days. However, investigators found a potential benefit of a lower oxygenation target in the 164 patients with suspected hypoxic–ischemic encephalopathy (relative risk, 0.73; 95% CI, 0.54 to 0.99). In the 332 patients with cardiac arrest in our trial, there was no clear between-group difference in 90-day mortality according to the randomized oxygenation targets, although firm conclusions cannot be drawn (Table S3).

The strengths of our trial include the variety of ICUs and countries involved and the pragmatic protocol that called for maintaining routine practice except for the oxygenation targets, while obtaining a clear between-group difference in Pao_2, Sao_2, and Fio_2 levels. Limitations must also be considered. The oxygenation targets that we used in our trial may have differed from standard of care in some countries. In a post hoc assessment, we found potential differences in the treatment effects among the individual ICUs (Fig. S4). We tested the two oxygen-therapy strategies by targeting intermittent measurement of the Pao_2; however, to account for the varying sampling schedules, all the patients had continuous monitoring of the peripheral oxygen saturation. Measurement of the Pao_2 may allow for more accurate maintenance of oxygenation targets than other methods, since the peripheral oxygen saturation can substantially differ from the Sao_2 under certain conditions[24,25] and may be less accurate in Black patients than in White patients.[26] However, targeting the Pao_2 is less feasible without placement of an arterial line and without the availability of point-of-care blood gas analysis. The use of standardized conversion tables for Fio_2 in open systems is another limitation, since the oxygen content in the lung varies with the patient's breathing patterns among other factors. Our evaluation of the between-group difference in values for Pao_2, Fio_2, and Sao_2 was limited by a diminishing number of patients in the ICU after the initial 14 to 21 days.

In a meta-analysis,[14] investigators reported the possibility that more liberal oxygen therapy in acutely ill adults may result in increased mortality.[14] However, an updated systematic review and meta-analysis with trial sequential analysis, including the ICU-ROX trial[16] among others, showed neither beneficial nor harmful effects of higher versus lower oxygenation strategies.[27] Although we found no differences in clinical outcomes between the two oxygenation groups in adults with

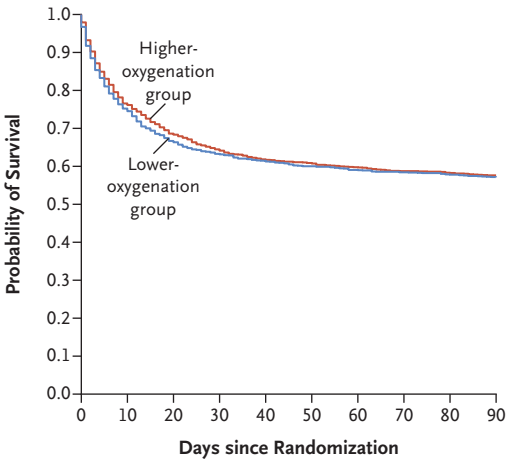


Figure 3. Kaplan–Meier Estimates of Survival.

Shown are the results of Kaplan–Meier analysis of data regarding survival, which were administratively censored at 90 days (adjusted hazard ratio, 1.04; 95% confidence interval, 0.93 to 1.16). The Cox proportional-hazards model was adjusted for the trial site and for the presence or absence of chronic obstructive pulmonary disease or active hematologic cancer.

acute hypoxemic respiratory failure, the results do not preclude the possibility of clinically important harm or benefit with the lower oxygenation strategy.

Thus, a lower oxygenation target did not result in lower mortality at 90 days than a higher-oxygenation target among patients in the ICU with acute hypoxemic respiratory failure.

Supported by a grant (4108-00011A) from Innovation Fund Denmark, by the Aalborg University Hospital, by grants (EMN-2017-00901 and EMN-2019-01055) from the Regions of Denmark, by a grant (25457) from the Obel Family Foundation, by the Danish Society of Anesthesiology and Intensive Care Medicine, and by the Intensive Care Symposium Hindsgavl.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

APPENDIX

The authors' full names and academic degrees are as follows: Olav L. Schjørring, M.D., Ph.D., Thomas L. Klitgaard, M.D., Anders Perner, M.D., Ph.D., Jørn Wetterslev, M.D., Ph.D., Theis Lange, Ph.D., Martin Siegemund, M.D., Minna Bäcklund, M.D., Ph.D., Frederik Keus, M.D., Jon H. Laake, M.D., Ph.D., Matthew Morgan, M.D., Ph.D., Katrin M. Thormar, M.D., Ph.D., Søren A. Rosborg, M.D., Ph.D., Jannie Bisgaard, M.D., Ph.D., Annette E.S. Erntgaard, M.D., Anne-Sofie H. Lynnerup, M.D., Rasmus L. Pedersen, M.D., Elena Crescioli, M.D., Theis C. Gielstrup, M.D., Meike T. Behzadi, M.D., Lone M. Poulsen, M.D., Stine Estrup, M.D., Ph.D., Jens P. Laigaard, M.S., Cheme Andersen, M.D., Camilla B. Mortensen, R.N., Björn A. Brand, M.D., Jonathan White, M.D., Ph.D., Inge-Lise Jarnvig, M.D., Ph.D., Morten H. Møller, M.D., Ph.D., Lars Quist, M.D., Ph.D., Morten H. Bestle, M.D., Ph.D., Martin Schønemann-Lund, M.D., Maj K. Kamper, M.D., Mathias Hindborg, M.D., Alexa Hollinger, M.D., Caroline E. Gebhard, M.D., Núria Zellweger, M.Sc., Christian S. Meyhoff, M.D., Ph.D., Mathias Hjort, M.D., Laura K. Bech, M.Sc., Thorbjørn Grøfte, M.D., Ph.D., Helle Bundgaard, M.D., Ph.D., Lars H.M. Østergaard, M.D., Maria A. Thyø, M.D., Thomas Hildebrandt, M.D., Bülent Uslu, M.D., Christoffer G. Sølling, M.D., Ph.D., Nette Møller-Nielsen, M.D., Anne C. Brøchner, M.D., Ph.D., Morten Borup, M.D., Ph.D., Marjatta Okkonen, M.D., Ph.D., Willem Dieperink, Ph.D., Ulf G. Pedersen, M.D., Anne S. Andreasen, M.D., Ph.D., Lone Buus, M.D., Tayyba N. Aslam, M.D., Robert R. Winding, M.D., Joerg C. Schefold, M.D., Stine B. Thorup, M.D., Susanne A. Iversen, M.D., Janus Engstrøm, B.Sc., Maj-Brit N. Kjær, R.N., and Bodil S. Rasmussen, M.D., Ph.D.

The authors' affiliations are as follows: the Department of Anesthesia and Intensive Care, Aalborg University Hospital, Aalborg University, Aalborg (O.L.S., T.L.K., S.A.R., J.B., A.E.S.E., A.-S.H.L., R.L.P., E.C., T.C.G., M.T.B., B.S.R.), Zealand University Hospital, Køge (L.M.P., S.E., J.P.L., C.A., C.B.M.), Rigshospitalet (A.P., B.A.B., J. Wetterslev, I.-L.J., M.H.M., L.Q., M.-B.N.K.) and the Section of Biostatistics, Department of Public Health (T.L.), University of Copenhagen, Bispebjerg, and Frederiksberg Hospital (C.S.M., M. Hjort, L.K.B.) and the Copenhagen Trial Unit (J. White, J.E.), Copenhagen, Nordsjaellands Hospital, Hillerod (M.H.B., M.S.-L., M.K.K., M. Hindborg), Randers Hospital, Randers (T.G., H.B.), Aarhus University Hospital, Skejby North, Aarhus (L.H.M.Ø., M.A.T.), Zealand University Hospital, Roskilde (T.H., B.U.), Viborg Hospital, Viborg (C.G.S., N.M.-N.), Kolding Hospital, Kolding (A.C.B., M. Bäcklund), Hvidovre Hospital, Hvidovre (U.G.P.), Herlev Hospital, Herlev (A.S.A.), Horsens Hospital, Horsens (L.B.), Herning Hospital, Herning (R.R.W.), Holbæk Hospital, Holbæk (S.B.T.), and Slagelse Hospital, Slagelse (S.A.I.) — all in Denmark; the Department of Intensive Care and Department of Clinical Research, University of Basel, Basel (M.S., A.H., C.E.G., N.Z.), and the Department of Intensive Care, Inselspital, University of Bern, Bern (J.C.S.) — both in Switzerland; the Department of Perioperative, Intensive Care and Pain Medicine, Helsinki University Hospital, Helsinki (M.B., M.O.); the Department of Critical Care, University Medical Center Groningen, Groningen, the Netherlands (F.K., W.D.); the Departments of Anesthesia and Intensive Care and of Clinical Research, Rikshospitalet, Oslo University Hospital, Oslo (J.H.L., T.N.A.); the Department of Intensive Care, Cardiff University Hospital of Wales, Cardiff, United Kingdom (M.M.); and the Department of Anesthesia and Intensive Care, Landspitali University Hospital, Reykjavik, Iceland (K.M.T.).

REFERENCES

1. Bellani G, Laffey JG, Pham T, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. JAMA 2016;315:788-800.

2. Helmerhorst HJF, Arts DL, Schultz MJ, et al. Metrics of arterial hyperoxia and associated outcomes in critical care. Crit Care Med 2017;45:187-95.

3. Schjørring OL, Jensen AKG, Nielsen CG, et al. Arterial oxygen tensions in mechanically ventilated ICU patients and mortality: a retrospective, multicentre, observational cohort study. Br J Anaesth 2020;124:420-9.

4. Claesson J, Freundlich M, Gunnarsson I, et al. Scandinavian clinical practice guideline on mechanical ventilation in adults with the acute respiratory distress syndrome. Acta Anaesthesiol Scand 2015; 59:286-97.

5. Fan E, Del Sorbo L, Goligher EC, et al. An official American Thoracic Society/European Society of Intensive Care Medicine/Society of Critical Care Medicine clinical practice guideline: mechanical ventilation in adult patients with acute respiratory distress syndrome. Am J Respir Crit Care Med 2017;195:1253-63.

6. O'Driscoll BR, Howard LS, Earis J, Mak V. BTS guideline for oxygen use in adults in healthcare and emergency settings. Thorax 2017;72:Suppl 1:ii1-ii90.

7. Beasley R, Chien J, Douglas J, et al. Thoracic Society of Australia and New Zealand oxygen guidelines for acute oxygen use in adults: 'swimming between the flags.' Respirology 2015;20:1182-91.

8. Panwar R, Hardie M, Bellomo R, et al. Conservative versus liberal oxygenation targets for mechanically ventilated patients: a pilot multicenter randomized controlled trial. Am J Respir Crit Care Med 2016;193:43-51.

9. Girardis M, Busani S, Damiani E, 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;316:1583-9.

10. The 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. N Engl J Med 2000;342:1301-8.

11. Brower RG, Lanken PN, MacIntyre N, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. N Engl J Med 2004;351:327-36.

12. Aggarwal NR, Brower RG, Hager DN, et al. Oxygen exposure resulting in arterial oxygen tensions above the protocol goal was associated with worse clinical outcomes in acute respiratory distress syndrome. Crit Care Med 2018;46:517-24.

13. Schjørring OL, Toft-Petersen AP, Kusk KH, et al. Intensive care doctors' prefer-

ences for arterial oxygen tension levels in mechanically ventilated patients. Acta Anaesthesiol Scand 2018;62:1443-51.
14. Chu DK, Kim LH-Y, Young PJ, 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; 391:1693-705.
15. Barrot L, Asfar P, Mauny F, et al. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. N Engl J Med 2020;382:999-1008.
16. Mackle D, Bellomo R, Bailey M, et al. Conservative oxygen therapy during mechanical ventilation in the ICU. N Engl J Med 2020;382:989-98.
17. Schjørring OL, Perner A, Wetterslev J, et al. Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) — protocol for a randomised clinical trial comparing a lower vs a higher oxygenation target in adults with acute hypoxaemic respiratory failure. Acta Anaesthesiol Scand 2019;63:956-65.
18. Schjørring OL, Klitgaard TL, Perner A, et al. The handling oxygenation targets in the intensive care unit (HOT-ICU) trial: detailed statistical analysis plan. Acta Anaesthesiol Scand 2020;64:847-56.
19. Fergusson D, Aaron SD, Guyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ 2002; 325:652-4.
20. Goodman SN. Toward evidence-based medical statistics. 2. The Bayes factor. Ann Intern Med 1999;130:1005-13.
21. Jakobsen JC, Tamborrino M, Winkel P, et al. Count data analysis in randomised clinical trials. J Biom Biostat 2015;6 (http://dx.doi.org/10.4172/2155-6180 .1000227).
22. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. Serial evaluation of the SOFA score to predict outcome in critically ill patients. JAMA 2001;286:1754-8.
23. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials — a practical guide with flowcharts. BMC Med Res Methodol 2017;17:162.
24. Ebmeier SJ, Barker M, Bacon M, et al. A two centre observational study of simultaneous pulse oximetry and arterial oxygen saturation recordings in intensive care unit patients. Anaesth Intensive Care 2018;46:297-303.
25. Schjørring OL, Rasmussen BS. The paramount parameter: arterial oxygen tension versus arterial oxygen saturation as target in trials on oxygenation in intensive care. Crit Care 2018;22:324.
26. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. N Engl J Med 2020;383:2477-8.
27. Barbateskovic M, Schjørring OL, Krauss SR, et al. Higher vs lower oxygenation strategies in acutely ill adults: a systematic review with meta-analysis and trial sequential analysis. Chest 2021;159: 154-73.

Appendix B. Paper II

Lower versus higher oxygenation targets in ICU patients with severe hypoxaemia: secondary Bayesian analyses of mortality and heterogeneous treatment effects in the HOT-ICU trial

Klitgaard TL, Schjørring OL, Lange T, Møller MH, Perner A, Rasmussen BS, Granholm A.

British Journal of Anaesthesia. 2021. October 18.

(Online ahead of print October 19, 2021)

doi: 10.1016/j.bja.2021.09.010

Link: sciencedirect.com/science/article/pii/S0007091221005821?via%3Dihub

BJA

British Journal of Anaesthesia, 128 (1): 55–64 (2022)

doi: 10.1016/j.bja.2021.09.010
Advance Access Publication Date: 19 October 2021
Critical Care

CRITICAL CARE

Lower *versus* higher oxygenation targets in critically ill patients with severe hypoxaemia: secondary Bayesian analysis to explore heterogeneous treatment effects in the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial

Thomas L. Klitgaard[1,2,3,*,†], Olav L. Schjørring[1,2,3,†], Theis Lange[3,4], Morten H. Møller[3,5], Anders Perner[3,5], Bodil S. Rasmussen[1,2,3,‡] and Anders Granholm[3,5]

[1]Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark, [2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark, [3]Collaboration for Research in Intensive Care, Copenhagen, Denmark, [4]Department of Public Health, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark and [5]Department of Intensive Care 4131, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

*Corresponding author. E-mail: tlk@rn.dk

[†]Coordinating investigators.
[‡]Sponsor and principal investigator of the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial.

Abstract

Background: In the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial, a lower (8 kPa) *vs* a higher (12 kPa) PaO_2 target did not affect mortality amongst critically ill adult patients. We used Bayesian statistics to evaluate any heterogeneity in the effect of oxygenation targets on mortality between different patient groups within the HOT-ICU trial.

Methods: We analysed 90-day all-cause mortality using adjusted Bayesian logistic regression models, and assessed heterogeneous treatment effects according to four selected baseline variables using both hierarchical models of subgroups and models with interactions on the continuous scales. Results are presented as mortality probability (%) and relative risk (RR) with 95% credibility intervals (CrI).

Results: All 2888 patients in the intention-to-treat cohort of the HOT-ICU trial were included. The adjusted 90-day mortality rates were 43.0% (CrI: 38.3–47.8%) and 42.3% (CrI: 37.7–47.1%) in the lower and higher oxygenation groups, respectively (RR 1.02 [CrI: 0.93–1.11]), with 36.5% probability of an RR <1.00. Analyses of heterogeneous treatment effects suggested a dose–response relationship between baseline norepinephrine dose and increased mortality with the lower oxygenation target, with 95% probability of increased mortality associated with the lower oxygenation target as norepinephrine doses increased.

Conclusions: A lower oxygenation target was unlikely to affect overall mortality amongst critically ill adult patients with acute hypoxaemic respiratory failure. However, our results suggest an increasing mortality risk for patients with a lower oxygen target as the baseline norepinephrine dose increases. These findings warrant additional investigation.

Clinical trial registration: NCT03174002.

Keywords: Bayesian analysis; heterogeneity of treatment effects; intensive care unit; oxygen therapy; respiratory insufficiency

Received: 9 June 2021; **Accepted:** 17 September 2021

Editor's key points

- Bayesian statistics can provide a valuable alternative perspective on clinical trial findings, particularly where knowing the most likely treatment effect can alter clinical practice even if this finding is not certain.
- The authors identified important differences in the effect of lower oxygenation targets between patient subgroups, which could be important in the care of critically ill adults.
- The possibility that critically ill patients in haemodynamic shock are more exposed to harm with lower oxygenation targets is important and should be investigated further in ongoing randomised trials.

Patients acutely admitted to the ICU with hypoxaemic respiratory failure are treated with supplemental oxygen. This treatment is believed to be life-saving, but the optimal target for oxygen therapy is not fully established. No firm conclusion on the benefits and harms of a lower *vs* a higher oxygenation target has been drawn for patients admitted to the ICU, as shown in a recently published systematic review.[1] This may be because of limited data, or to a large degree of heterogeneity in published trials.

In the Normal Oxygenation Versus Hyperoxia in the Intensive Care Unit (OXYGEN-ICU) trial, a lower oxygenation strategy resulted in noticeably reduced ICU mortality compared with a higher oxygenation strategy in a mixed cohort of ICU patients (8.6 percentage points difference; 95% confidence interval [CI]: 1.7–15.0%), but the trial was stopped at an unplanned interim analysis after an earthquake.[2] The Liberal Oxygenation Versus Conservative Oxygenation in ARDS ($LOCO_2$) trial suggested benefit from a higher oxygenation strategy compared with a lower oxygenation strategy because of a reduced mortality at both 28 days (7.8 percentage points difference; 95% CI: −4.8 to 20.6) and 90 days post-randomisation (14.0 percentage points difference; 95% CI: 0.7–27.2%).[3] However, this trial was also stopped early, as an unplanned interim analysis found observations of intestinal ischaemia, an unplanned secondary outcome, in the lower oxygenation group, but not in the higher oxygenation group. The Intensive Care Unit Randomized Trial Comparing Two Approaches to Oxygen Therapy (ICU-ROX) trial found no differences in 28-day ventilator-free days (−0.3 days absolute difference; 95% CI: −2.1 to 1.6 days) or in 90-day mortality (odds ratio [OR] 1.10; 95% CI: 0.84–1.44) between a lower and a higher oxygenation strategy.[4] In the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial, adult patients with acute hypoxaemic respiratory failure in the ICU were randomised to an arterial partial pressure of oxygen (PaO_2) of 8 kPa (lower target) or 12 kPa (higher target) during ICU admission.[5] At 90 days, 42.9% of patients in the lower oxygenation group had died and 42.4% in the higher oxygenation group, resulting in an adjusted relative risk (RR) of 1.02 (95% CI: 0.94–1.11) in the primary frequentist analysis. Comparable results were found in the conventional subgroup analyses.[5] However, heterogeneous treatment effects may still be present.[6–8]

Bayesian statistical methods allow for detailed probabilistic quantifications of effect sizes, and integration of prior knowledge allows for nuanced sensitivity analyses of the intervention effects. Such methods have previously been used in several large-scale trials to complement the conventional frequentist analysis[9–12] or as the primary statistical framework.[13–15] In this prospective Bayesian analysis of the HOT-ICU trial,[16] our aim was to provide a probabilistic evaluation of the effects of a lower oxygenation target *vs* a higher oxygenation target on 90-day all-cause mortality, to assess the probabilities of a number of pre-specified effect sizes, including effects larger than the *a priori* hypothesised 20% relative reduction in mortality,[17,18] and to explore the presence of heterogeneous treatment effects on mortality based on pre-specified baseline variables.

Methods

This secondary Bayesian analysis of the HOT-ICU trial was conducted in accordance with a protocol and statistical analysis plan published before randomisation of the last patient,[16] and prepared according to recent recommendations.[6,8,19,20] It was guided by the same principles as the Bayesian analysis of heterogeneous treatment effects in the Stress Ulcer Prophylaxis in the Intensive Care Unit (SUP-ICU) trial.[12,21] The results are reported according to the Reporting of Bayes Used in clinical STudies (ROBUST) guideline,[22] and this paper has been prepared in agreement with the Strengthening the Reporting of Observational Studies in Epidemiology statement.[23]

HOT-ICU trial

The HOT-ICU trial was an investigator-initiated international, pragmatic, parallel-group, stratified, randomised trial (RCT), which enrolled patients from June 20, 2017 to August 3, 2020. Adult patients (≥18 yr), acutely admitted to the ICU with hypoxaemic respiratory failure, receiving a fraction of inspired oxygen (FiO_2) of at least 0.50 in a closed system (invasive or noninvasive mechanical ventilation or mask/helmet CPAP) or at least oxygen 10 L min^{-1} in an open system, had an arterial line, and were expected to receive supplemental oxygen for at least 24 h in the ICU were included. Patients were randomised 1:1 to the lower oxygenation target or the higher oxygenation target, which was applied during the entire ICU stay, including readmissions, for up to 90 days. Additional details on the HOT-ICU trial, including exclusion criteria, approvals, and variable definitions, are available in the Supplementary Appendix and elsewhere.[5,17,18]

Outcome measure

The primary outcome measure was 90-day all-cause mortality.

Statistical analysis

All statistical analyses were performed using R version 4.0.4 (R Core Team, R Foundation for Statistical Computing, Vienna, Austria) and Stan[24] through the *brms* R package,[25,26] with additional details available in the Supplementary Appendix. We used Bayesian logistic regression models that incorporated prior distributions expressing pre-existing beliefs of effect sizes and their uncertainties in combination with data from the trial at hand. The models combined this to inform posterior distributions of the variables of interest.[27] Posterior distributions were summarised using median values and percentile-based 95% credibility intervals (CrI) that may be interpreted as the 95% most probable values, conditional on

the priors, models and data.[28] The full posterior distributions were presented graphically, supplemented with probabilities of pre-specified and additional effect sizes.[16] Results were presented as posterior adjusted risk ratios (RRs) and risk differences (RDs), and adjusted event probabilities in each group (used to calculate RRs and RDs), calculated by setting adjustment variables to their most common value, as specified in the protocol.[16] We also present the results on the underlying odds ratio (OR) scale to facilitate comparison with other studies that may have reported on this scale. Relative risk and OR <1, and RD <0 favoured the lower oxygenation target; RR and OR >1, and RD >0 favoured the higher oxygenation target.

Priors

For the primary analysis of the intervention effect, we used weakly informative priors centred on no difference (OR of 1=RR of 1) and including a large range containing all plausible effect sizes (ORs with 95% probability between 0.14 and 7.10). We thus expected the trial data to dominate the posterior probability distributions because of the large sample size of the HOT-ICU trial. Two pre-specified sensitivity analyses were conducted: (i) using evidence-based priors informed by an updated random-effects meta-analysis of previous RCTs, and (ii) using sceptic priors centred on no difference and sceptical of larger effect sizes, as described in the protocol.[16] Full details on priors are presented in the Supplementary Appendix and in the protocol.[16]

Subgroup-based heterogeneity of treatment effect analyses

We assessed the presence of heterogeneous treatment effects using four different subgrouping schemes based on selected baseline variables:

(i) Sequential Organ Failure Assessment (SOFA) score as a marker of organ dysfunction[29]
(ii) PaO_2:FiO_2 ratio as a marker of severity of hypoxaemic respiratory failure with additional adjustment for the type of oxygen supplementation system at baseline (closed or open), with closed system being the reference
(iii) Highest continuously infused dose of norepinephrine during the 24 h before randomisation
(iv) Latest plasma lactate concentration before randomisation

Five quintile-based subgroups were created of each variable ensuring that all patients with identical values were in the same groups. We used hierarchical Bayesian logistic regression models with partial pooling adjusted for the stratification variables (chronic obstructive pulmonary disease, haematological malignancy, and site) to calculate subgroup results.[26,30] Results were presented using the effect measures outlined previously. Additional information on parameter definitions is available in the Supplementary Appendix and elsewhere.[5]

Continuous heterogeneity of treatment effect analyses

We assessed the potential interactions of the allocation to the lower oxygenation target with the four baseline characteristics of interest for 90-day all-cause mortality on the continuous scale using Bayesian logistic regression models. All models were adjusted for the stratification variables mentioned previously. Additional adjustment for type of oxygen supplementation system (open or closed) at baseline was performed when assessing PaO_2:FiO_2 ratio. Results are presented using conditional effects plots with ORs and 95% CrI for interactions, and probabilities for interaction ORs <1 (negative interaction) and >1 (positive interaction). The conditional effects plots illustrate the predicted probabilities of an outcome dependent on the variables of interest (treatment, the baseline variable, and their interaction), with all other variables kept constant at their reference values (adjustment variables set to their most common values).

Missing data and technical model details

We planned *a priori* to use complete case analysis if missingness for all variables in an analysis was less than 5% and multiple imputation otherwise.[16] For all Bayesian models, we used four chains with 5000 warm-up and 5000 post-warm-up draws per chain, yielding 20 000 post-warm-up draws in all. For additional details on handling of missing data and model diagnostics, see the Supplementary Appendix and the protocol.[16]

Results

We included 2888 of the 2928 patients (98.6%) randomised in the HOT-ICU trial, equivalent to the full intention-to-treat cohort.[5] Baseline characteristics of the trial cohort are presented in Table 1. Additional characteristics of all subgroups according to quintiles and stratified according to treatment allocation are presented in Supplementary Tables 1a–4b. Diagnostics for all statistical models were acceptable.

Bayesian analysis of 90-day all-cause mortality

The adjusted RR for mortality was 1.02 (95% CrI: 0.93–1.11), with 63.5% probability of an RR >1.00. The probability of an RR <0.80, equivalent to the 20% *a priori* hypothesised relative mortality reduction,[17] or more was <0.01%. We observed similar low probabilities (<2%) of such effect sizes across all subgroups, except for low plasma lactate concentrations (Supplementary Table 6). The full posterior probability distribution for 90-day all-cause mortality is presented in Fig. 1 (RD and OR distributions are presented in Supplementary Fig. 1a and b). Probabilities for mortality along with RRs and RDs for the trial cohort are presented in Table 2 (ORs are available in Supplementary Table 5).

Subgroup-based heterogeneity of treatment effect analyses

A substantial number of patients did not receive norepinephrine at baseline; these patients were all included in the same subgroup, which is thus larger than the remaining four quartile-based subgroups. The apparent overlap amongst PaO_2:FiO_2 ratio-based subgroup limits is attributable to rounding (Table 2).

For increasing baseline doses of norepinephrine, we found increasing risk for 90-day all-cause mortality, indicating benefit of the higher oxygenation target: from RR 0.99 (95% CrI: 0.87–1.11) in the lowest dosage group (all 0.00 mM) to RR 1.08 (95% CrI: 0.95–1.33) in the highest dosage group (0.40–2.40 mM). This potential dose–response relationship was not found in any of the other baseline variable subgrouping schemes. Posterior probabilities for mortality and the estimates of RRs and RDs in the four sets of subgroups are

Table 1 Baseline characteristics for all patients. Baseline characteristics for the trial cohort stratified by oxygenation target allocation. Numerical values are presented as medians with inter-quartile ranges (IQRs) and categorical variables as numbers (*n*) and percentages (%). FiO_2, fraction of inspired oxygen; PaO_2, partial pressure of arterial oxygen; SaO_2, saturation of arterial oxygen; SOFA, Sequential Organ Failure Assessment. Additional baseline characteristics are available in the primary trial publication.[5] *The PaO_2:FiO_2 ratio was missing in five patients in the lower oxygenation group and in seven patients in the higher oxygenation group. †Plasma lactate concentration was missing in eight patients in the lower oxygenation group and in 11 patients in the higher oxygenation group. ‡The aggregated SOFA score ranges from 0 to 24, with sub-score from 0 to 4 for six organ systems (respiration, coagulation, liver, cardiovascular, CNS, and renal), with higher scores indicating higher degrees of organ failure. The SOFA score was missing in 44 patients in the lower oxygenation group and in 45 patients in the higher oxygenation group because of one or more missing sub-scores of the SOFA score.

Variable	Lower target, *n*=1441	Higher target, *n*=1447
Median age (IQR, yr)	70 (61–77)	70 (60–77)
Male sex, *n* (%)	916 (63.6)	939 (64.9)
Type of admission, *n* (%)		
Medical	1238 (85.9)	1233 (85.2)
Elective surgical	18 (1.3)	21 (1.5)
Emergency surgical	185 (12.8)	193 (13.3)
Chronic obstructive pulmonary disease	277 (19.2)	285 (19.7)
Active haematological cancer	81 (5.6)	86 (5.9)
Oxygen supplementation in a closed system, *n* (%)	1024 (71.1)	1038 (71.7)
Invasive mechanical ventilation, *n* (%)	826 (57.3)	863 (59.6)
Noninvasive ventilation or CPAP, *n* (%)	198 (13.7)	175 (12.1)
Oxygen supplementation in an open system, *n* (%)	417 (28.9)	409 (28.3)
Median PaO_2 (IQR, kPa)	10.3 (8.7–12.6)	10.3 (8.7–12.3)
Median FiO_2 (IQR)	0.70 (0.55–0.90)	0.70 (0.58–0.85)
Median PaO_2:FiO_2 ratio (IQR)*		
In all systems	15.8 (11.8–21.0)	15.7 (12.0–20.5)
In closed systems	16.5 (12.2–21.7)	16.5 (12.6–21.4)
In open systems	14.1 (10.9–18.4)	13.9 (10.7–18.0)
Median lactate concentration (IQR, mM)†	1.8 (1.1–3.2)	1.7 (1.1–3.1)
Any use of vasopressors, *n* (%)	793 (55.0)	785 (54.3)
Median highest dose of norepinephrine (IQR, μg kg^{-1} min^{-1})	0.20 (0.10–0.40)	0.21 (0.10–0.40)
Median SOFA score (IQR)‡	8 (5–10)	8 (5–10)

presented in Table 2 (ORs are presented in Supplementary Table 5). The posterior probability distribution plots of the RRs for mortality in the subgroups are presented in Fig. 2 (RD and OR distributions are presented in Supplementary Fig. 4a and b). The posterior probabilities for different RRs for all four sets of subgroups are presented in Supplementary Table 6. Comparisons of treatment effects in the subgroups are presented in Supplementary Tables 11–14.

Continuous heterogeneity of treatment effect analyses

We found a 95% probability of a positive interaction between increasing baseline norepinephrine dose and the lower oxygenation target on mortality (i.e. unfavourable effects of a lower oxygenation target with increasing dose of norepinephrine at baseline). For increasing baseline lactate concentrations, the probability of a positive interaction with the lower oxygenation target on mortality was 86% (i.e. potential increased mortality risk of the lower oxygenation target for patients with higher concentrations of lactate). The probabilities of positive interactions (i.e. potential increased mortality risks) between the lower oxygenation target and the remaining baseline variables were 65% for increasing baseline SOFA scores (i.e. higher degree of organ failure) and 76% for decreasing baseline PaO_2:FiO_2 ratios (i.e. greater severity of respiratory failure). Conditional effect plots showing the estimated interactions between treatment allocation and baseline variables on mortality on the continuous scale are presented in Fig. 3.

Sensitivity analyses

The results of the sensitivity analyses using evidence-based and sceptic priors were largely consistent with the findings of the primary analysis (Supplementary Table 7; Supplementary Figs 2a–3c and 5a–7b).

Missing data

No imputation of missing data was performed, as missingness was <5% for all variables of interest included in any analysis.[18] For additional details on missing data, see the Supplementary Appendix and elsewhere.[5]

Discussion

In this prospective, secondary analysis of treatment effects in the HOT-ICU trial, the risk of death within 90 days for patients treated with a lower oxygenation target was with 95% probability between RR 0.93 and 1.11. Given these data, larger effect sizes are improbable. Our analyses suggested heterogeneous treatment effects when considering the interaction between the lower oxygenation target and baseline norepinephrine dose, suggesting that in patients with higher degrees of shock (measured as higher administered doses of continuously infused norepinephrine), a lower oxygenation strategy may be harmful. This effect was consistent across a series of models. A similar trend was identified in the continuous model assessing plasma lactate concentrations at baseline, but without indications of the same relation in the subgroup-based heterogeneity analyses, and thus with no clear support for a dose–response relationship. Caution must be used when interpreting these findings, as the effect was only suggested in one of the two models. We found no strong suggestions of heterogeneous treatment effects according to SOFA scores or PaO_2:FiO_2 ratios at baseline.

The results of the Bayesian analysis of the 90-day all-cause mortality in this study are consistent with the primary frequentist analysis of the HOT-ICU trial,[5] the ICU-ROX trial,[4] and the latest meta-analysis conducted before the publication of

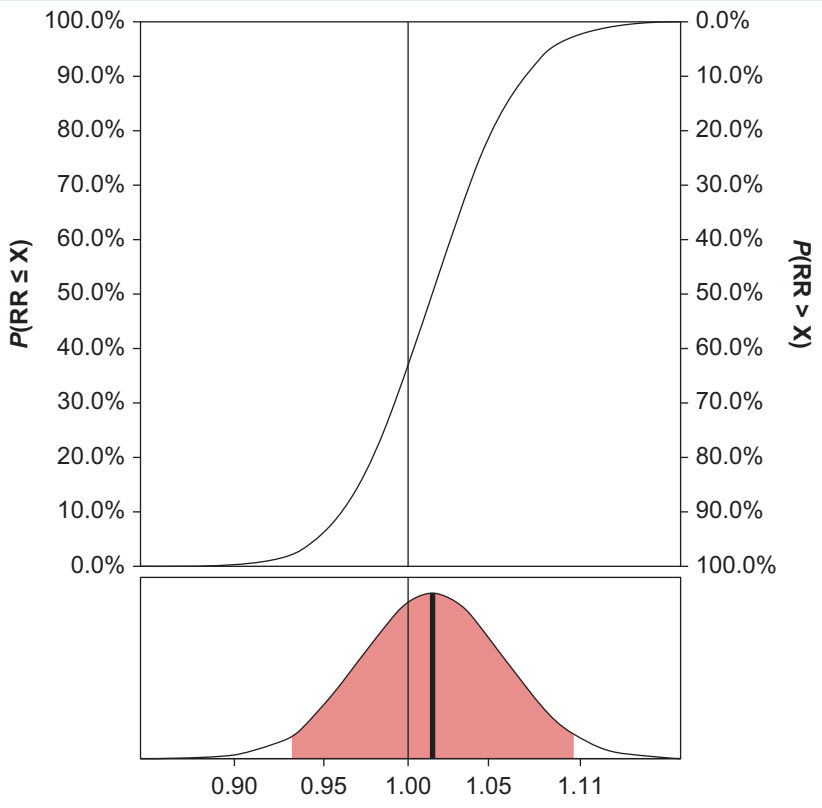


Fig 1. Posterior probability distribution for the adjusted relative risk (RR) for 90-day all-cause mortality in the primary analysis using *weakly informative* priors. Upper part: cumulative posterior probability distribution for the adjusted RR. *P*(RR ≤ X) is the probability that the RR is smaller or equal to any given value specified on the X-axis, being 'X'; *P*(RR > X) is the probability that the RR is larger than any given value specified on the X-axis, being 'X'. An RR <1 indicates benefit from the lower oxygenation target; an RR >1 indicates benefit of the higher oxygenation target. Lower part: full posterior probability distribution; full vertical line=median value; coloured area=95% credibility interval.

the HOT-ICU trial.[1] In contrast, the OXYGEN-ICU trial demonstrated benefit from a conservative oxygenation strategy,[2] whilst the LOCO$_2$ trial found potential benefit of a more liberal oxygenation strategy.[3] However, given the substantially smaller sizes of the OXYGEN-ICU and LOCO$_2$ trials (n=480 and 205, respectively) compared with the HOT-ICU (n=2928) and the ICU-ROX (n=1000) trials, and the fact that both were stopped after unplanned interim analyses, the findings of these trials may be attributable to chance. Also, the inclusion criteria of the trials differ substantially, as the ICU-ROX[4] and LOCO$_2$[3] trials included only invasively mechanically ventilated patients, whereas the OXYGEN-ICU[2] and HOT-ICU[5] trials included patients on both open and closed oxygen supplementation systems. Additionally, when considering baseline PaO_2:FiO_2 ratios, patients presented with substantially more severe respiratory failure in the LOCO$_2$[3] and HOT-ICU[5] trials compared with the ICU-ROX[4] trial. These aspects may impede direct comparison of the results. Although larger effect sizes for mortality in the broad population of adult patients in the ICU with acute severe hypoxaemic respiratory failure seem improbable, smaller effects may also be of importance. Even a 2% absolute reduction in mortality would result in 2000 lives saved for every 100 000 patients treated with supplemental oxygen. The ongoing MEGA-ROX[31] and UK-ROX[32] trials are designed to assess absolute risk reductions for mortality of 1.5 and 2.5 percentage points, respectively, comparing a lower *vs* a higher oxygenation target. Effect sizes of such magnitudes cannot be excluded based on our results.

None of the aforementioned trials[2–4] have considered the presence of heterogeneous treatment effects in a comparable manner to the one presented here. However, in a subgroup of patients with sepsis in the ICU-ROX trial, point estimates of treatment effects indicated harm of a lower oxygenation strategy, although this was not statistically significant.[33] Similar was found in the subgroup of patients with shock at baseline in the HOT-ICU trial.[5] On the contrary, the OXYGEN-ICU trial found reduced occurrence of shock when using a conservative oxygenation strategy compared with a more liberal oxygenation strategy.[2]

The strengths and limitations from the HOT-ICU trial are all carried over to this study.[5] The most important strengths are the size of the trial, the pragmatic design, high external validity (35 ICUs in seven countries), and the clear separation in the oxygenation parameters between the intervention groups.[5] Also, the protocol for this study was published before randomisation of the last patient in the HOT-ICU trial.[16] Further, our results were consistent in the sensitivity analyses using different priors, and we evaluated the presence of

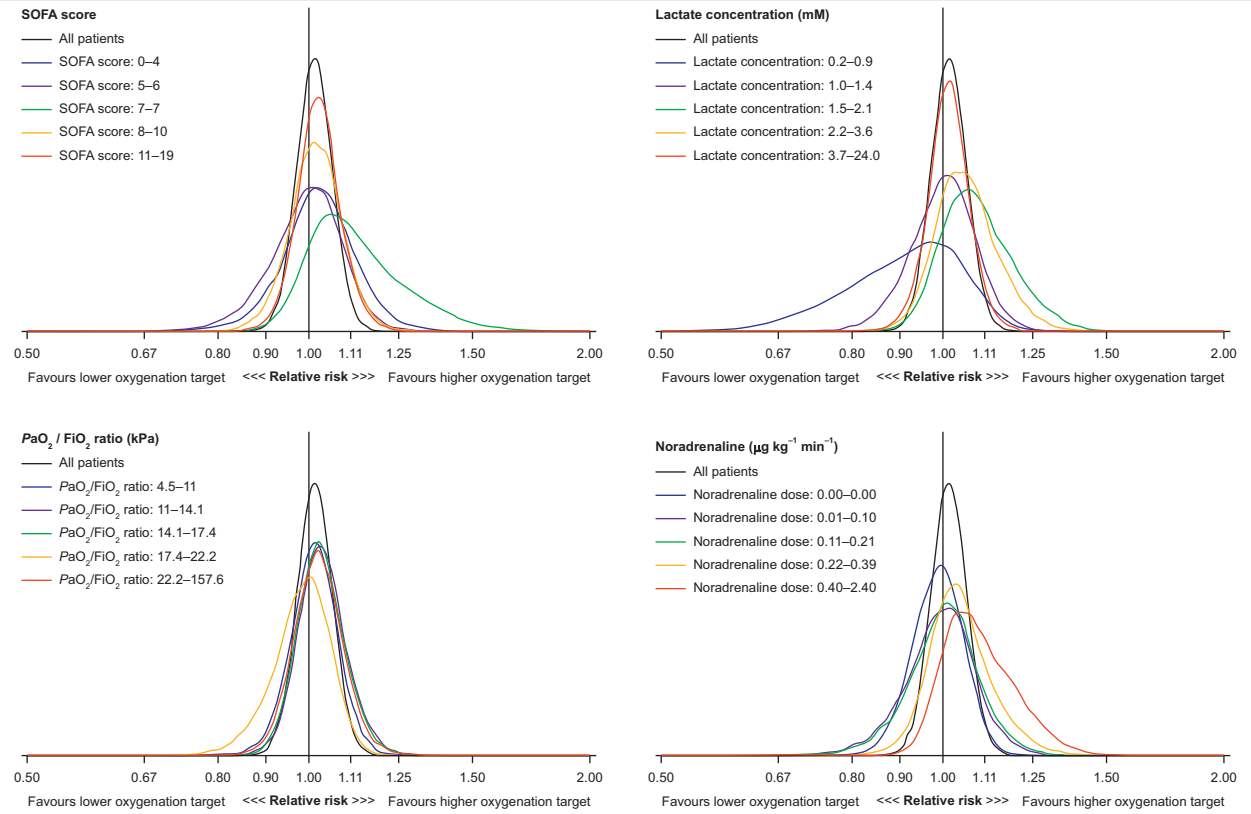


Fig 2. Posterior probability distributions of the adjusted relative risks (RRs) of the treatment effect on 90-day all-cause mortality according to the four pre-specified baseline variables in the primary analysis using weakly informative priors. The posterior probability distributions of RRs in each subgroup from the subgroup-based models are displayed together with the posterior distribution from the corresponding analysis of all patients not considering subgroups. An RR <1 indicates benefit from the lower oxygenation target; an RR >1 indicates benefit of the higher oxygenation target. PaO_2:$FiO_2$$FiO_2$, ratio of partial pressure of arterial oxygen to fraction of inspired oxygen; SOFA, Sequential Organ Failure Assessment.

heterogeneity of treatment effects both in subgroups and on the continuous scale, which may ease interpretation of our finding and serves as a consistency check. The limitations of this study are mainly related to the heterogeneity of treatment effect analyses. We chose the variables of interest based on availability and of the following reasons:[16] the SOFA score is independently associated with mortality,[34] and assessment of heterogeneity of treatment effects according to the risk of the outcome is recommended.[8] Based on clinical rationale, different degrees of hypoxaemic respiratory failure may benefit from different levels of oxygenation; plasma lactate concentration and norepinephrine dose both serve as markers of shock, which, in turn, is associated with increased mortality.[35] A dedicated prediction model for mortality would have been preferable, but this was not available. Also, other variables, or combinations of such, could have provided additional information on the potential heterogeneity with different oxygenation targets. As some subgroups may contain few events, this may lead to imprecision. Yet, this effect is to some extent mitigated by shrinkage and partial pooling in the hierarchical models.[26,30] As the categorisation of the continuous baseline variables into quintile-based subgroups was data driven, cut-offs did not follow established conventions (e.g. in relation to the PaO_2:FiO_2 ratio), limiting the generalisability of the results. However, this was chosen to ensure that all subgroups were of adequate and similar sizes. In the analyses on the continuous scale, we assumed a linear relationship (on the log-OR scale) between the variables of interest and mortality, including the interaction term. For the sake of simplicity and to limit the risk of spurious findings and overfitting because of the use of multiple and increasingly flexible models, no other models to predict this relationship were applied. Lastly, secondary analyses and subgroup analyses should always be cautiously interpreted. Despite the analyses being preplanned and the benefits of the Bayesian methods, the risks of spurious findings are not eliminated. All results from this study should consequently be regarded as hypothesis generating only.

In conclusion, the RR for 90-day all-cause mortality, when comparing a lower oxygenation target with a higher oxygenation target in adult patients in the ICU with acute hypoxaemic respiratory failure, was between 0.93 and 1.11 with 95% probability. Based on this, larger effect sizes are highly improbable. Our findings also suggest potentially important

Table 2 Summarised effect measures for 90-day all-cause mortality. Adjusted posterior event probabilities, relative risks (RRs), and risk differences (RDs) for 90-day all-cause mortality in the primary analysis using weakly informative priors. CrI, credibility interval; SOFA, Sequential Organ Failure Assessment; PaO_2:FiO_2, ratio of partial pressure of arterial oxygen to fraction of inspired oxygen ratio; *n*, number of patients in each group (after excluding patients with missing data for one or more variables included in the analyses). RR <1 and RD <0 favour the lower target; RR >1 and RD >0 favour the higher target. *The SOFA score ranges from 0 to 24, with sub-score from 0 to 4 for six organ systems (respiration, coagulation, liver, cardiovascular, CNS, and renal), with higher aggregated scores indicating higher degrees of organ failure. †PaO_2:FiO_2 ratio: lower scores indicate more severe pulmonary dysfunction.

Group	*n*	Event probability, lower target (%)	Event probability, higher target (%)	RR	RD (%)
All patients	2888	43.0 (95% CrI: 38.3–47.8)	42.3 (95% CrI: 37.7–47.1)	1.02 (95% CrI: 0.93–1.11)	0.6 (95% CrI: −3.0 to 4.3)
SOFA score (baseline)*	2799				
0–4	486	32.5 (95% CrI: 26.5–39.1)	31.7 (95% CrI: 25.8–38.3)	1.03 (95% CrI: 0.85–1.23)	0.8 (95% CrI: −5.3 to 6.5)
5–6	501	35.5 (95% CrI: 29.3–42.1)	35.7 (95% CrI: 29.5–42.6)	1.00 (95% CrI: 0.81–1.16)	0.0 (95% CrI: −7.2 to 5.3)
7–7	352	37.6 (95% CrI: 30.5–45.7)	33.6 (95% CrI: 26.3–41.0)	1.10 (95% CrI: 0.94–1.48)	3.4 (95% CrI: −2.3 to 13.6)
8–10	881	42.1 (95% CrI: 36.3–48.0)	41.4 (95% CrI: 35.8–47.3)	1.02 (95% CrI: 0.89–1.15)	0.7 (95% CrI: −4.7 to 5.9)
11–19	579	57.2 (95% CrI: 50.7–63.5)	55.8 (95% CrI: 49.4–62.1)	1.02 (95% CrI: 0.92–1.15)	1.4 (95% CrI: −4.6 to 7.7)
Lactate concentration (baseline, mM)	2869				
0.2–0.9	501	23.1 (95% CrI: 17.5–29.2)	25.4 (95% CrI: 19.9–32.0)	0.92 (95% CrI: 0.66–1.14)	−1.9 (95% CrI: −10.0 to 3.1)
1.0–1.4	631	38.1 (95% CrI: 32.1–44.6)	38.0 (95% CrI: 32.0–44.6)	1.00 (95% CrI: 0.85–1.16)	0.2 (95% CrI: −6.3 to 5.8)
1.5–2.1	577	42.0 (95% CrI: 35.5–49.2)	38.7 (95% CrI: 32.2–45.3)	1.08 (95% CrI: 0.93–1.32)	3.1 (95% CrI: −2.7 to 11.1)
2.2–3.6	576	45.0 (95% CrI: 38.8–51.7)	42.5 (95% CrI: 36.0–49.1)	1.06 (95% CrI: 0.92–1.25)	2.3 (95% CrI: −3.5 to 9.6)
3.7–24.0	584	61.7 (95% CrI: 55.0–67.9)	60.8 (95% CrI: 54.2–67.0)	1.01 (95% CrI: 0.91–1.13)	0.9 (95% CrI: −5.5 to 7.1)
Norepinephrine dose (baseline, $\mu g\ kg^{-1}\ min^{-1}$)	2888				
0.00–0.00	1373	38.1 (95% CrI: 33.0–43.5)	38.6 (95% CrI: 33.4–44.0)	0.99 (95% CrI: 0.87–1.11)	−0.4 (95% CrI: −5.3 to 4.0)
0.01–0.10	366	39.8 (95% CrI: 32.5–47.3)	40.1 (95% CrI: 33.2–47.3)	1.00 (95% CrI: 0.82–1.17)	−0.1 (95% CrI: −7.8 to 6.3)
0.11–0.21	372	39.5 (95% CrI: 32.4–47.0)	39.5 (95% CrI: 32.6–46.4)	1.01 (95% CrI: 0.83–1.19)	0.2 (95% CrI: −7.3 to 6.9)
0.22–0.39	348	50.0 (95% CrI: 42.4–57.6)	47.8 (95% CrI: 40.4–55.5)	1.04 (95% CrI: 0.91–1.24)	1.8 (95% CrI: −4.9 to 10.4)
0.40–2.40	429	52.4 (95% CrI: 45.3–60.2)	48.0 (95% CrI: 40.9–55.2)	1.08 (95% CrI: 0.95–1.33)	3.9 (95% CrI: −2.5 to 14.0)
PaO_2:FiO_2 ratio (baseline, kPa)†	2876				
4.5–11.0	565	46.0 (95% CrI: 39.8–52.4)	45.3 (95% CrI: 39.6–51.5)	1.02 (95% CrI: 0.90–1.14)	0.7 (95% CrI: −4.8 to 5.8)
11.0–14.1	584	46.6 (95% CrI: 40.4–53.3)	45.1 (95% CrI: 39.5–51.1)	1.03 (95% CrI: 0.92–1.17)	1.4 (95% CrI: −3.6 to 7.4)
14.1–17.4	574	46.6 (95% CrI: 40.5–53.1)	45.2 (95% CrI: 39.5–51.3)	1.03 (95% CrI: 0.92–1.16)	1.3 (95% CrI: −3.7 to 7.0)
17.4–22.2	577	41.6 (95% CrI: 34.8–48.3)	42.4 (95% CrI: 36.1–48.4)	0.99 (95% CrI: 0.84–1.11)	−0.5 (95% CrI: −7.2 to 4.5)
22.2–157.6	576	44.0 (95% CrI: 37.7–50.4)	43.0 (95% CrI: 36.9–48.8)	1.02 (95% CrI: 0.91–1.16)	1.0 (95% CrI: −4.2 to 6.5)

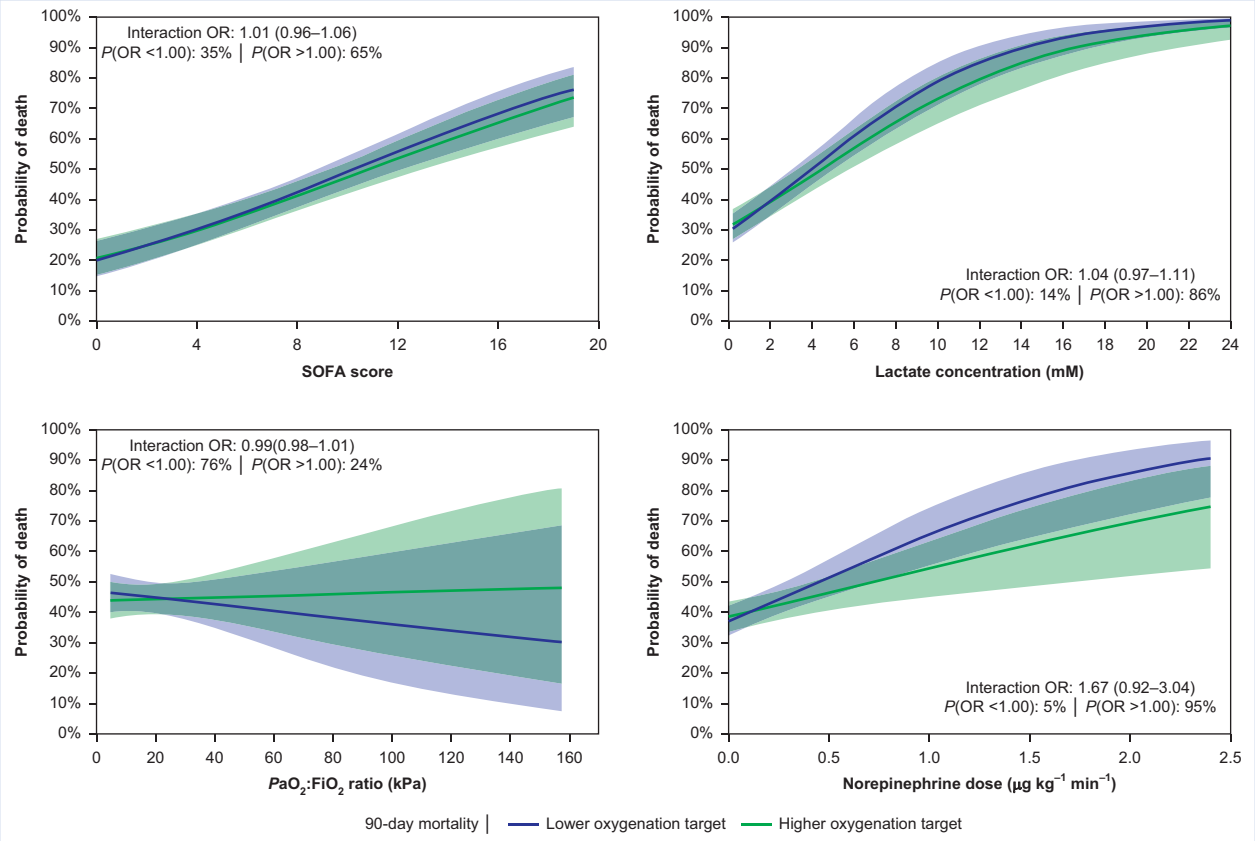


Fig 3. Conditional effects plots for 90-day all-cause mortality, using weakly informative priors. These plots illustrate the estimated interactions between treatment allocation and 90-day all-cause mortality on the continuous scale. The levels of the individual variables of interest are plotted on the X-axes; the probabilities of mortality are plotted on the Y-axes. Within each subplot, the odds ratio (OR) with 95% credibility interval for the interaction effect between the lower oxygenation target and the baseline variable assessed is presented. The posterior probabilities that the interaction OR is <1.00 (negative interaction) or >1.00 (positive interaction) are also presented. PaO_2:F_{IO2}, ratio of partial pressure of arterial oxygen to fraction of inspired oxygen; SOFA, Sequential Organ Failure Assessment. In total, 95% of patients had a PaO_2:FiO_2 ratio <35.5 kPa.

heterogeneity in treatment effects in terms of baseline norepinephrine dose as an index of haemodynamic shock. This increasing probability of death for patients treated with lower oxygenation targets as norepinephrine dose increases requires further investigation.

Authors' contributions

Study conception: AG, TLK, OLS, MHM, AP, BSR
Statistical analysis plan and protocol: all authors
Involved in the Conducting of the Handling Oxygenation Targets in the Intensive Care Unit trial: all authors
Analyses: TLK, AG
Writing of first draft: TLK
Critical revision: all authors
Approval of paper: all authors

Acknowledgements

The authors would like to express their gratitude to all involved in the Handling Oxygenation Targets in the Intensive Care Unit trial: research staff and investigators, clinical staff, patients, and their relatives.

Declarations of interest

The Department of Intensive Care at Rigshospitalet has received funding for other projects from the Novo Nordisk Foundation, Pfizer, and Fresenius Kabi.

Funding

Innovation Fund Denmark (4108-00011A); Aalborg University Hospital; Regions of Denmark (EMN-2017-00901 and EMN-2019-01055); Obel Family Foundation (25457); Danish Society of Anaesthesiology and Intensive Care Medicine; Intensive Care Symposium Hindsgavl.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2021.09.010.

References

1. Barbateskovic M, Schjørring OL, Krauss SR, et al. Higher vs lower oxygenation strategies in acutely ill adults. *Chest* 2020; **159**: 154–73
2. Girardis M, Busani S, Damiani E, 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**: 1583–9
3. Barrot L, Asfar P, Mauny F, et al. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. *N Engl J Med* 2020; **382**: 999
4. The ICU-ROX investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group. Conservative oxygen therapy during mechanical ventilation in the ICU. *N Engl J Med* 2020; **382**: 989–98
5. Schjørring OL, Klitgaard TL, Perner A, et al. Lower or higher oxygenation targets for acute hypoxemic respiratory failure. *N Engl J Med* 2021; **384**: 1301–11
6. Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC. Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med* 2015; **192**: 1045–51
7. Young PJ. Effect of oxygen therapy on mortality in the ICU. *N Engl J Med* 2021; **384**: 1361–3
8. Kent DM, Steyerberg E, Van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ* 2018; **363**: k4245
9. Goligher EC, Tomlinson G, Hajage D, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome and posterior probability of mortality benefit in a post hoc Bayesian analysis of a randomized clinical trial. *JAMA* 2018; **320**: 2251–9
10. Zampieri FG, Costa EL, Iwashyna TJ, et al. Heterogeneous effects of alveolar recruitment in acute respiratory distress syndrome: a machine learning reanalysis of the Alveolar Recruitment for Acute Respiratory Distress Syndrome trial. *Br J Anaesth* 2019; **123**: 88–95
11. Zampieri FG, Damiani LP, Bakker J, et al. Effects of a resuscitation strategy targeting peripheral perfusion status versus serum lactate levels among patients with septic shock. A Bayesian reanalysis of the ANDROMEDA-SHOCK trial. *Am J Respir Crit Care Med* 2020; **201**: 423–9
12. Granholm A, Marker S, Krag M, et al. Heterogeneity of treatment effect of prophylactic pantoprazole in adult ICU patients: a post hoc analysis of the SUP-ICU trial. *Intensive Care Med* 2020; **46**: 717–26
13. The REMAP-CAP Investigators. Interleukin-6 receptor antagonists in critically ill patients with Covid-19. *N Engl J Med* 2021; **384**: 1491–502
14. The Writing Committee for the REMAP-CAP Investigators. Effect of hydrocortisone on mortality and organ support in patients with severe COVID-19: the REMAP-CAP COVID-19 corticosteroid domain randomized clinical trial. *JAMA* 2020; **324**: 1317–29
15. Angus DC, Berry S, Lewis RJ, et al. The REMAP-CAP (randomized embedded multifactorial adaptive platform for community-acquired pneumonia) study rationale and design. *Ann Am Thorac Soc* 2020; **17**: 879–91
16. Klitgaard TL, Schjørring OL, Lange T, et al. Bayesian and heterogeneity of treatment effect analyses of the HOT-ICU trial—a secondary analysis protocol. *Acta Anaesthesiol Scand* 2020; **9**: 1376–81
17. Schjørring OL, Perner A, Wetterslev J, et al. Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU)—protocol for a randomised clinical trial comparing a lower vs a higher oxygenation target in adults with acute hypoxaemic respiratory failure. *Acta Anaesthesiol Scand* 2019; **63**: 956–65
18. Schjørring OL, Klitgaard TL, Perner A, et al. The handling oxygenation targets in the intensive care unit (HOT-ICU) trial: detailed statistical analysis plan. *Acta Anaesthesiol Scand* 2020; **64**: 847–56
19. Ferreira D, Barthoulot M, Pottecher J, Torp KD, Diemunsch P, Meyer N. A consensus checklist to help clinicians interpret clinical trial results analysed by Bayesian methods. *Br J Anaesth* 2020; **125**: 208–15
20. Ferreira D, Barthoulot M, Pottecher J, Torp KD, Diemunsch P, Meyer N. Theory and practical use of Bayesian methods in interpreting clinical trial data: a narrative review. *Br J Anaesth* 2020; **125**: 1–7
21. Granholm A, Marker S, Krag M, et al. Heterogeneity of treatment effect of stress ulcer prophylaxis in ICU patients: a secondary analysis protocol. *Acta Anaesthesiol Scand* 2019; **63**: 1251–6
22. Sung L, Hayden J, Greenberg ML, Koren G, Feldman BM, Tomlinson GA. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J Clin Epidemiol* 2005; **58**: 261–8
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008; **61**: 344–9
24. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. *J Stat Softw* 2017; **76**: 1–31
25. Bürkner PC. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw* 2017; **80**: 1–28
26. Bürkner PC. Advanced Bayesian multilevel modeling with the R package brms. *R J* 2018; **10**: 395–411
27. Kruschke J. *Doing bayesian data analysis—a tutorial with R, JAGS, and stan*. 2nd Edn. Cambridge, MA: Academic Press; 2014
28. Ryan EG, Harrison EM, Pearse RM, Gates S. Perioperative haemodynamic therapy for major gastrointestinal surgery: the effect of a Bayesian approach to interpreting the findings of a randomised controlled trial. *BMJ Open* 2019; **9**, e024256
29. Vincent J-L, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med* 1996; **22**: 707–10
30. McGlothlin AE, Viele K. Bayesian hierarchical models. *JAMA* 2018; **320**: 2365–6
31. ANZICS. MEGA-ROX trial (ANZICS reg.no. CTG1920-01) 2021. Available from https://www.anzics.com.au/current-active-endorsed-research/mega-rox/ (accessed April 14 2021).
32. Intensive Care National Audit & Research Centre. UK-ROX (ICNARC project number: NIHR130508) 2021. Available from https://www.icnarc.org/Our-Research/Studies/Uk-Rox (accessed April 14 2021).
33. Young P, Mackle D, Bellomo R, et al. Conservative oxygen therapy for mechanically ventilated adults with sepsis: a post hoc analysis of data from the intensive care unit randomized trial comparing two approaches to

oxygen therapy (ICU-ROX). *Intensive Care Med* 2020; **46**: 17–26

34. Granholm A, Møller MH, Krag M, Perner A, Hjortrup PB. Predictive performance of the Simplified Acute Physiology Score (SAPS) II and the initial Sequential Organ Failure Assessment (SOFA) score in acutely ill intensive care patients: post-hoc analyses of the SUP-ICU inception cohort study. *PLoS One* 2016; **11**, e0168948
35. Singer M, Deutschman CS, Seymour C, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016; **315**: 801–10

Handling editor: Rupert Pearse

Appendix C. Paper III

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

Klitgaard TL, Schjørring OL, Nielsen FM, Meyhoff CS, Perner A, Wetterslev J, Rasmussen BS, Barbateskovic M.

Submitted to the *Cochrane Database of Systematic Reviews* on January 12, 2022. Art. No.: CD012631.
doi: 10.1002/14651858.CD012631.pub3

Cochrane Library

Cochrane Database of Systematic Reviews

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

Lass Klitgaard T, Schjørring OL, Mølgaard Nielsen F, Meyhoff CS, Perner A, Wetterslev J, Barbateskovic M, Rasmussen BS

Lass Klitgaard T, Schjørring OL, Mølgaard Nielsen F, Meyhoff CS, Perner A, Wetterslev J, Barbateskovic M, Rasmussen BS.
Higher versus lower fractions of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit.
Cochrane Database of Systematic Reviews TBD, Issue TBD. Art. No.: CD012631.
DOI: 10.1002/14651858.CD012631.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

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

Thomas Lass Klitgaard[1,2,3], Olav L Schjørring[1,2,3], Frederik Mølgaard Nielsen[1,2,3], Christian S Meyhoff[4], Anders Perner[3,5], Jørn Wetterslev[3,6], Marija Barbateskovic[6], Bodil S Rasmussen[1,2,3]

[1]Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark. [2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark. [3]Centre for Research in Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [4]Department of Anaesthesia and Intensive Care, Bispebjerg and Frederiksberg Hospital, University of Copenhagen, Copenhagen, Denmark. [5]Department of Intensive Care, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [6]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Contact: Thomas Lass Klitgaard, tlk@rn.dk.

Editorial group: Cochrane Emergency and Critical Care Group.
Publication status and date: New search for studies and content updated (conclusions changed), published in Issue , .

Citation: Lass Klitgaard T, Schjørring OL, Mølgaard Nielsen F, Meyhoff CS, Perner A, Wetterslev J, Barbateskovic M, Rasmussen BS. Higher versus lower fractions of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit. *Cochrane Database of Systematic Reviews* TBD, Issue TBD. Art. No.: CD012631. DOI: 10.1002/14651858.CD012631.pub2.

ABSTRACT

Background

This is an updated review concerning 'higher versus lower fractions of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit'.

Oxygen is the most widely used medical drug, and is provided to the vast majority of patients admitted to the intensive care unit (ICU) to prevent global and organ hypoxia. Oxygen supplementation has been administered liberally, resulting in a proportion of patients with hyperoxaemia. This has been associated with increased mortality and morbidity in some settings, but not in others. Thus far, only limited data have been available to inform clinical practice guidelines, and the optimum oxygenation target for adults admitted to the ICU remains undetermined. Even though solid evidence of benefits remains scarce, the provision of supplemental oxygen is still recommended in practice guidelines. However, one must strive to achieve the optimum balance between potentially harmful effects of hyperoxaemia and potential beneficial effects of supplemental oxygen.

Objectives

To update the assessment of benefits and harms of higher versus lower fractions of inspired oxygen (FiO_2) or targets of arterial oxygenation for adults admitted to the ICU.

Search methods

We searched CENTRAL, MEDLINE, Embase, Science Citation Index Expanded, BIOSIS Previews and LILACS. We searched for ongoing or unpublished trials in clinical trials registers, and scanned the reference lists and citations of included studies. Literature searches for this updated review were conducted in April 2021.

Selection criteria

We included randomized controlled trials (RCTs) that compared higher versus lower FiO_2 or targets of arterial oxygenation (partial pressure of oxygen (PaO_2), peripheral or arterial oxygen saturation (SpO_2 or SaO_2)) for adults admitted to the ICU. We included trials irrespective of publication type, publication status, and language.

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

Data collection and analysis

Four review authors independently screened the references identified in the literature searches and extracted the data. Our primary outcomes were all-cause mortality, the proportion of participants with one or more serious adverse events (SAEs), and quality-of-life. We analysed all outcomes at maximum follow-up. Only three trials reported the proportion of participants with one or more SAEs as according to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP) criteria. However, most trials reported on events categorised by us as SAEs. We therefore conducted two post-hoc analyses of the effect of higher versus lower oxygenation strategies using 1) the single SAE with the highest reported proportion in each trial and 2) the cumulated proportion of participants with an SAE in each trial. One trial reported on quality-of-life.

Secondary outcomes were occurrence of lung injury, myocardial infarction, stroke, and sepsis.

No trial reported on lung injury as a composite outcome, but four trials reported on the occurrence of acute respiratory distress syndrome (ARDS) and four on pneumonia. We updated the two post-hoc meta-analyses of the effect of higher versus lower oxygenation strategies using 1) the single lung injury event with the highest reported proportion in each trial and 2) the cumulated 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 using the Risk of Bias 2 tool. We used the GRADEpro tool to assess the overall certainty of the evidence. We also evaluated the risk of publication bias for outcomes reported by more than ten trials.

Main results

We included 16 RCTs (6486 participants), of which 14 reported relevant outcomes for this review (6349 participants). For all-cause mortality, eight trials were judged to be at overall low risk of bias, and five at overall high risk of bias. For the reported SAEs, eight trials were judged to be at overall low risk of bias, and six at overall high risk of bias. The one trial reporting on quality-of-life was judged to be at overall low risk of bias.

Meta-analysis of all trials regardless of risk of bias indicated no evidence of a difference from higher or lower oxygenation strategies at maximum follow-up with regard to mortality (risk ratio (RR) 1.01, 95% CI 0.94 to 1.10; I^2 = 9%; 13 trials; 5973 participants; very low-certainty evidence), or occurrence of SAEs: proportion of patients with one or more SAE RR 1.05 (95% CI 0.98 to 1.13; I^2 = 27%; 3744 participants; 3 trials; low certainty evidence), the highest proportion of specific SAEs in each trial RR 1.00 (95% CI 0.95 to 1.06; I^2 = 38%; 6031 participants; 14 trials). However, trial sequential analyses could reject a relative risk increase or reduction of 10% for mortality and 20% for SAEs. Given the low-certainty of evidence it is necessary to interpret these findings with caution.

Only one of the included trials reported data on quality of life at any time point, indicating no evidence of a difference between higher or lower oxygenation strategies.

Meta-analysis of all trials indicated no evidence of a difference between higher or lower oxygenation strategies on the occurrence of lung injuries at maximum follow-up (the highest reported proportion of lung injury RR 1.06, 95% CI 0.82 to 1.36; I^2 = 0%; 1942 participants; 7 trials; very low-certainty evidence).

Meta-analysis of all trials indicated harm from higher oxygenation strategies as compared with lower on the occurrence of sepsis at maximum follow-up. Meta-analysis indicated no differences with regard to the occurrences of myocardial infarction or stroke.

Authors' conclusions

In adult ICU patients, there is still uncertainty about the effects of higher versus lower oxygenation strategies on all-cause mortality, SAEs, quality of life, lung injuries, myocardial infarction, stroke, and sepsis at maximum follow-up due to low to very low certainty evidence.

PLAIN LANGUAGE SUMMARY

Supplemental oxygen for adults admitted to the intensive care unit

Review question

We set out to update the assessment on 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 have a high risk of dying. Oxygen supplementation, or therapy, is provided to most adult ICU patients 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 high oxygen levels, known as hyperoxia. Despite a lack of robust evidence of effectiveness, supplemental oxygen administration has been widely recommended in international clinical practice guidelines. However, newer guidelines recommend against high oxygen levels as some, but not all, trials have indicated a link between hyperoxaemia and an increased risk of dying. The potential benefits of supplemental oxygen must be weighed against the potentially harmful effects of hyperoxia.

Trial characteristics

We identified 16 randomized controlled trials where participants were randomly allocated to either a higher or a lower oxygen supplementation strategy involving 6486 participants up to September 2021. Fourteen of the trials (6349 participants) provided findings on the number of deaths, serious adverse events, quality of life, or lung injuries at any time-point following oxygen therapy in the ICU. The occurrence of lung injury was measured according to participants developing acute respiratory distress syndrome or pneumonia. Nine trials included adults admitted to the ICU due to various serious health conditions; three trials included medically ill patients only; and two included surgical patients only. Two trials assessed adults with traumatic brain injury; one trial assessed adults resuscitated from out-of-hospital cardiac arrest; and one trial assessed adults with stroke. In eight trials, all participants received invasive mechanical ventilation via a tube inserted into the trachea. Six trials involved patients both mechanically ventilated and not. Two trials involved adults receiving any non-invasive oxygen administration. The use of more oxygen was compared with less oxygen in all trials, but the levels of such differed greatly.

Oxygen therapy was provided for a variety of periods of time, ranging from one hour to the entire hospital admission (up to 90 days).

Key results

After this update, we are still uncertain about the effects of higher versus lower oxygen supplementation strategies as our findings are based on low-certainty evidence.

We did not find evidence for a beneficial effect of higher compared with lower oxygen supplemental strategies for adult ICU patients, neither on the risk of death (13 trials; 5973 participant), the occurrence of one or more serious adverse event (3 trials; 3744 participants), the highest proportion of serious adverse events (14 trials, 6031 participants), the quality of life (1 trial, 499 participants), the risk of lung injury (7 trials; 1942 participant), the risk of myocardial infarction (3 trials, 3368 participants), the risk of stroke (4 trials, 4476 participants), nor the risk of sepsis (2 trials, 646 participants). The evidence is, however, still very uncertain.

Certainty of the evidence

The number of participants enrolled in the trials was too small to permit a definitive judgement about the interventions effect sizes on the outcomes in this review. 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 this was supplied.

SUMMARY OF FINDINGS

Summary of findings 1. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU — trials at overall low risk of bias only

Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU — excluding trials at overall high risk of bias

Patient or population: adults admitted to the ICU
Setting: excluding trials at overall high risk of bias; conducted in ICU departments in Europe (n = 5); China (n = 1); Australia, New Zealand (n = 1); Australia, New Zealand, France (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 (trials)	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: range 1 month to 6 months	Study population		**RR 0.99** (0.91 to 1.09)	4945 (8 RCTs)	⨁⨁◯◯ Low[1]	-
	404 per 1000	**400 per 1000** (368 to 440)				
Proportion of participants with 1 or more serious adverse events according to International Conference on Harmonisation Good Clinical Practice (ICH-GCP): range 3 to 180 days	Study population		**RR 1.07** (0.99 to 1.15)	3344 (2 RCTs)	⨁⨁◯◯ Low[2]	Two sensitivity analyses were performed; estimated highest proportion reported and estimated cumulated number of events. Meta-analysis from the analysis of the highest proportion of serious adversed events reported which addresses the lowest possible proportion of participants with 1 or more serious adverse events showed RR 1.00 (95% CI 0.94 to 1.06; I^2 = 42%; 4945 participants; 8 RCTs). Meta-analysis from the analysis of cumulating all reported serious adverse events which address the highest possible reported proportion of participants
	413 per 1000	442 per 1000 (409 to 475)				

						with 1 or more serious adverse events showed RR 1.02 (95% CI 0.98 to 1.06; I^2 = 0%; 4212 participants; 8 RCTs).
Lung injury diagnosed after randomization (composite outcome): range 4 to 23 days	Study population		**RR 1.16** (0.74 to 1.81)	424 (3 RCTs)	⊕⊖⊖⊖ Very low[4]	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: 1 trial, 201 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 1.16 (95% CI 0.74 to 1.81; I^2 = 0%; 424 participants; 3 RCTs).
	142 per 1000	**148 per 1000** (108 to 205)				
Myocardial infarction diagnosed after randomization: range 90 days	Study population		**RR 0.57** (0.24 to 1.35)	2910 (1 RCT)	⊕⊖⊖⊖ Very low[5]	Meta-analysis was not conducted, as only 1 trial reported on myocardial infarction and was judged to be at overall low risk of bias.
	10 per 1000	**5 per 1000** (2 to 13)				
Stroke diagnosed after randomization: range 1 to 3 months	Study population		**RR 1.04** (0.59 to 1.83)	3111 (2 RCTs)	⊕⊖⊖⊖ Very low[6]	-
	15 per 1000	**15 per 1000** (9 to 27)				
Sepsis diagnosed after randomization: range 90 days	Study population		**RR 1.68** (0.84 to 3.34)	201 (1 RCT)	⊕⊖⊖⊖ Very low[7]	Meta-analysis was not conducted, as only 1 trial reported on sepsis and was judged to be at overall low risk of bias.
	111 per 1000	**187 per 1000** (93 to 371)				

*__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 two levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials. Required information size (RIS) is 1,878 participants. RIS = optimal information size (OIS) when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[2]Downgraded two levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials. RIS is 1,804 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[3]Downgraded two levels: one level because we cannot reject inconsistency due to the inclusion of only one trial; and one level due to imprecision because RIS was not reached 500 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[4]Downgraded three levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials; and one level due to imprecision because RIS was not reached (8,509 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[5]Downgraded three levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials; and one level due to imprecision because RIS was not reached (141,612 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[6]Downgraded three levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials; and one level due to imprecision because RIS 93,883 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.
[7]Downgraded three levels: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials; and one level due to imprecision because RIS was not reached (128,595 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

Summary of findings 2. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU — all included trials

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

Patient or population: adults admitted to the ICU
Setting: trials were conducted in ICU departments in Europe (n = 8); China (n = 2); Iran (n = 2); Australia, 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 (trials)	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				

	gets of arterial oxygenation	**gets of arterial oxygenation**				
All-cause mortality: range 1 month to 6 months	Study population		**RR 1.01** (0.94 to 1.10)	5973 (13 RCTs)	⊕⊖⊖⊖ Very low[1]	-
	383 per 1000	**386 per 1000** (360 to 421)				
Proportion of participants with 1 or more serious adverse events according to International Conference on Harmonisation Good Clinical Practice (ICH-GCP): range 3 to 180 days	Study population		**RR 1.05** (0.98 to 1.13)	3744 (3 RCTs)	⊕⊖⊖⊖ Very low[2]	Two sensitivity analyses were performed; estimated highest proportion reported and estimated cumulated number of events. Meta-analysis from the analysis of the highest proportion of serious adversed events reported which addresses the lowest possible proportion of participants with 1 or more serious adverse events showed RR 1.00 (95% CI 0.95 to 1.06; I^2 = 38%; 6031 participants; 14 RCTs). Meta-analysis from the analysis of cumulating all reported serious adverse events which address the highest possible reported proportion of participants with 1 or more serious adverse events showed RR 1.03 (95% CI 1.00 to 1.06; I^2 = 67%; 6053 participants; 14 RCTs).
	415 per 1000	**442 per 1000** (409 to 475)				
Quality of life (any measure): range 180 days	Study population		Not estimable	499 (1 RCT)	⊕⊖⊖⊖ Very low[3]	Meta-analysis was not conducted, as only 1 trial reported on quality of life. Quality of life was reported as EQ-VAS scores.
	67.6 (± 22.4)	**70.1 (± 22.0)**				
Lung injury diagnosed after randomization (composite outcome): range 4 to 23 days	Study population		**RR 1.06** (0.82 to 1.36)	1942 (7 RCTs)	⊕⊖⊖⊖ Very low[4]	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: 4 trials, 862 participants; pneumonia: 4 trials, 1145 participants.
	143 per 1000	**154 per 1000** (119 to 198)				

						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 1.02 (95% CI 0.80 to 1.31; I^2 = 0%; 1942 participants; 7 RCTs). Meta-analysis from the analysis of ARDS showed RR 0.86 (95% CI 0.43 to 1.69; I^2 = 0%; 862; 4 RCTs). Meta-analysis from the analysis of pneumonia showed RR 1.08 (95% CI 0.82 to 1.41; I^2 = 0%; 1145 participants; 4 RCTs).
Myocardial infarction diagnosed after randomization: range up to 90 days	Study population 17 per 1000	**10 per 1000** (4 to 24	**RR 0.59** (0.25 to 1.38)	3368 (3 RCTs)	⊕⊖⊖⊖ Very low[5]	-
Stroke diagnosed after randomization: range 1 to 3 months	Study population 11 per 1000	**12 per 1000** (7 to 21)	**RR 1.12** (0.65 to 1.92)	4476 (4 RCTs)	⊕⊖⊖⊖ Very low[6]	-
Sepsis diagnosed after randomization: range 28 to 60 days	Study population 69 per 1000	**125 per 1000** (77 to 203)	**RR 1.81** (1.11 to 2.95)	646 (2 RCTs)	⊕⊖⊖⊖ Very low[7]	-

*__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 8 of 13 trials were judged to be at overall low risk of bias; two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials. Required information size (RIS) is 2,045 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 2 of 3 trials were judged to be at overall low risk of bias; two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials. Required information size (RIS) is 1,804 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

[3]Downgraded three levels: one level because we cannot reject inconsistency due to the inclusion of only one trial; one level due to risk of bias; and one level due to imprecision because RIS (500 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, as only 3 of 8 trials were judged to be at overall low risk of bias: two levels due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials and due to differences in inclusion criteria between trials. RIS is 8,438 participants. RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

[5]Downgraded three levels: one level because of risk of bias, as only 1 of 3 trials was judged to be at overall low risk of bias; one level due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level due to imprecision because RIS was not reached (82,653 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

[6]Downgraded three levels: one level because of risk of bias, as only 2 of 4 trials were judged to be at overall low risk of bias; one level due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level due to imprecision because RIS was not reached (117,747 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

[7]Downgraded three levels: one level because of risk of bias, as only 1 of 2 trials was judged to be at overall low risk of bias; one level due to indirectness because of differences in inspiratory oxygen fraction and target of arterial oxygenation in the experimental and control groups between trials; and one level due to imprecision because RIS was not reached (19,482 participants). RIS = OIS when $I^2 = 0$ and alpha is adjusted for multiple outcomes.

BACKGROUND

Description of the condition

In healthy individuals, the normal range for the partial pressure of arterial oxygen (PaO_2) at sea level is 10.7 kPa (80 mmHg) to 13.3 kPa (100 mmHg) (Kratz 2004), with a general decrease with age (Crapo 1999).

Patients admitted to the intensive care unit (ICU) are very frequently treated with supplemental oxygen to prevent or treat hypoxaemia and ultimately hypoxia. Hypoxaemia refers to lack of oxygen in the blood and is usually defined in terms of PaO_2 or arterial oxygen saturation of haemoglobin (SaO_2), 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). However, 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). Conversely, hyperoxia and hyperoxaemia refers to above normal levels of oxygen content in the body's tissues and blood, respectively. As with hypoxia and hypoxaemia, no clear definition of hyperoxia and hyperoxaemia exists, but with a suggested threshold of PaO_2 above 120 mmHg (O'Driscoll 2017). The peripheral oxygen saturation (SpO_2) measured by pulse oximetry is routinely used as a non-invasive surrogate for SaO_2. Currently, oxygenation targets below the normal range and even defined as hypoxaemic, targeting PaO_2 between 55 mmHg and 80 mmHg or SpO_2 between 88% and 95%, are employed in adults who are mechanically ventilated with acute respiratory distress syndrome (ARDS) in the ICU (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, right-to-left blood shunting, 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 a 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; Schjørring 2020).

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 FiO_2 has 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 (Stub 2015; 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 adult surgical patients (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 especially during mechanical ventilation. 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).

Oxygen administration has typically been liberal and have resulted in hyperoxaemia or hyperoxia in the lungs (de Graaff 2011; de Jonge 2008; Itagaki 2015; Kraft 2018; Panwar 2013; Rachmale 2012; Suzuki 2013; Zhang 2016; Schjørring 2020). Some observational studies and randomised trials have indicated an association between hyperoxaemia and mortality (Dahl 2015; Helmerhorst 2017a; Kilgannon 2010; Meyhoff 2012; Zhang 2016; Palmer 2019; Schjørring 2020), whilst other studies have not (Bellomo 2011; Eastwood 2012; Kraft 2018; Raj 2013; Young 2012), possibly because adults who receive excessive oxygen supplementation 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).

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). A recent systematic review on acutely ill patients found no evidence of a difference in mortality or serious adverse events when comparing the use of higher versus lower oxygenation strategies (Barbateskovic 2021 (a)) in contrast to a previous review of similar design (Chu 2018).

Although the possible adverse effects of hyperoxaemia are known, prevention of hypoxia through hyperoxaemia seems to be prioritised (Pannu 2016). The ideal oxygenation target 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).

Oxygen is a common intervention in adults admitted to the 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. This is an update of a Cochrane Review (Barbateskovic 2019).

OBJECTIVES

To update the assessment on 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 trials on any adult patients aged 18 years or older admitted to the ICU. We only included participants if they were admitted to the ICU when randomization was performed.

Types of interventions

We included trials having a clear differentiation of participants randomized to either a high or a low 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.

Experimental group: adults receiving a high oxygenation strategy administered by any device, the aim of which was to ensure adequate oxygenation through 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 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 FiO_2 or oxygenation. 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

We chose the following measures as outcomes.

Primary outcomes

1. All-cause mortality at maximum follow-up.
2. Proportion of participants with one or more serious adverse events (SAE), defined as a dichotomous outcome according to participants having at least one serious adverse event or none at maximum follow-up. We defined a serious adverse event as any untoward medical occurrence that: resulted in death; was life-threatening; required hospitalisation or prolongation of existing hospitalisation; resulted in persistent or significant disability; or jeopardized the participant, according to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP) (ICH-GCP 1997). We considered all other adverse events as non-serious (ICH-GCP 1997). We performed two additional analyses being the highest proportion reported of any specific SAE and the cumulated number of SAEs, on the proportion of participants with one or more SAE. As a secondary analysis, we analysed each SAE separately.
3. Quality of life (any valid scale such as the 36-item Short Form Health Survey (SF-36)) at maximum follow-up.

Secondary outcomes

1. Lung injury diagnosed after randomization (composite outcome) at maximum follow-up. This composite outcome was defined as either: ARDS (as defined by the Berlin criteria (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. Myocardial infarction diagnosed after randomization at maximum follow-up (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 maximum follow-up (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. Sepsis diagnosed after randomization at maximum follow-up (defined as sepsis plus sepsis-induced organ dysfunction or tissue hypoperfusion (Dellinger 2013), or as defined by trialists).

Search methods for identification of studies

We searched for studies as described in Cochrane Handbook of Systematic reviews of Intervention Chapter 4 (Lefebvre 2021). 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.

Electronic searches

We searched the following databases:

1. Cochrane Central Register of Controlled Trials (CENTRAL) (Issue 11, 2020 (Appendix 1);
2. MEDLINE (Ovid, 1946 to 20 April 2021) (Appendix 2);
3. Embase (Ovid, 1974 to 20 April 2021) (Appendix 3);
4. Web of Science/BIOSIS Previews (1969 to 20 April 2021) (Appendix 4);
5. Latin American and Caribbean Health Science Information database (LILACS) (1982 to 20 April 2021) (Appendix 5).

CINAHL was searched for the primary edition of this review (Barbateskovic 2019), but not for the updated version due to access restrictions.

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 two authors independently and in pair, searched for ongoing and unpublished trials in the following trial registers:

1. US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov) (searched 25 August 2021);
2. World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) (www.who.int/ictrp/en/) (searched 25 August 2021);
3. EU Clinical Trials Register (www.clinicaltrialsregister.eu/) (searched 25 August 2021);
4. Australian New Zealand Clinical Trials Registry (ANZCTR) (www.anzctr.org.au/) (searched 25 August 2021).

See Appendix 6 for search strategy.

We searched for systematic reviews in Epistemonikos (www.epistemonikos.org) .
Backward and forward citation searches for all included trials was performed using Web of Science.
We also checked for retractions using Retraction Watch database (retractiondatabase.org).
We used Covidence to deduplicate the references before screening the search result.
The searches were developed and run by the authors and peer reviewed by the Cochrane Emergency and Critical Care Information Specialist.
We contacted trial authors and experts in the field for additional information.

Data collection and analysis

We used the following methods for data collection and data analyses. Any discrepancies between the primary review and this updated version are described in detail in the following sections, and any changes made from the protocol are summarised in Differences between protocol and review.

Selection of studies

Four review authors (TLK, OLS, FMN, or MB), independently and in pair, 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. Disagreements were resolved by consensus or by consulting another review author (OLS or MB) when necessary.

Data extraction and management

Four authors (TLK, OLS, FMN, or MB) independently and in pair, extracted predefined data of the included trials using a data collection form that was specifically designed and piloted by the review team (Appendix 7). 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 (OLS or MB) when necessary.

Assessment of risk of bias in included studies

For this updated review we assessed risk of bias according to the latest *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021) using the 'Risk of Bias 2 tool' (RoB 2) (Higgins 2016; Sterne 2019), employing the criteria described in Appendix 8. Two review authors (TLK or FMN) independently assessed the methodological quality of each included trial and outcome, as defined by the design of the trial and reporting. Any disagreements were resolved by discussion or consultation with a third author (OLS or MB).

We assessed the following risk of bias domains for all included trials and outcomes: 1) risk of bias arising from the randomization process; 2) risk of bias due to deviations from the intended interventions (effect of assignment to intervention); 3) risk of bias due to missing outcome data; 4) risk of bias in measurement of the outcome; and 5) risk of bias in selection of the reported result.

Each domain was judged as being at "low risk of bias", "some concerns", or "high risk of bias". RCTs with "low risk of bias" in all domains were classified as being at overall "low risk of bias". RCTs with one domain judged to be at "some concerns", but no domain judged to be at "high risk of bias", where classified as being at overall "some concerns" of risk of bias. RCTs were classified as being at overall "high risk of bias" if at least one domain was judged as being at "high risk of bias". However, if a trial was judged to be at "some concerns" due to risk of bias for multiple domains, it may have been judged as being at overall "high risk of bias" if the assessors judged that the multiple concerns amounted to a serious risk of bias (Higgins 2016; Sterne 2019).

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 2016; Sterne 2019). This was also done for trials judged to be at overall 'low risk of bias' only.

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 due to updating with new trials 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 standardised 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

If trials were identified employing three different oxygenation targets, we would combine 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 with the control group. If only one of the experimental groups fulfilled the minimum difference to the control, this group was compared 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 with 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 2020).

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 2021). 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 visually assessed funnel plots for signs of asymmetry if an analysis included 10 or more trials (Higgins 2021; Jakobsen 2014a).

We tested asymmetry within dichotomous outcomes using the Harbord test (Harbord 2006), and for continuous outcomes using the asymmetry test (Egger 1997). We also used the 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 2021; Jakobsen 2014a), including TSA and calculation of Bayes factors. We performed meta-analyses 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 Web (RevMan Web, version 3.6.0, The Cochrane Collaboration, 28 June 2021, available at revman.cochrane.org) and the TSA software version 0.9 CTU to analyse data (Review Manager 2020; TSA 2011) and the STATA software version 16 (STATA 2019).

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 if the two models produced comparable point estimates.

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 trials, 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 anaesthesia related 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 2011; 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) for dichotomous outcomes and minimal important difference for continuous outcomes, 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 harm 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 predefined estimations of the anticipated intervention effect 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, i.e the expected intervention effect would equal the difference from no effect (RR = 1) and the confidence limit closest to 1. As a post-hoc sensitivity analysis we performed TSA using an anticipated RRR or RRR of 10%. See Differences between protocol and review.

No TSA-plot or TSA CI was presented if the information size for any outcome was less than 5%.

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 (equal to 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 oxygenation, the result is that more oxygen is delivered, and the oxygenation parameterswill be elevated in both strategies. However, we recognise that, especially in adults with ARDS, there are individuals where it would be extremely difficult to reach a predefined target of oxygenation 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 overall risk of bias:
   a. overall low risk og bias
   b. overall some concern
   c. overall high risk of bias
2. According to different types of oxygen interventions:
   a. oxygenation target measured using either PaO_2 or SaO_2 or SpO_2 (as defined by trialists)
   b. oxygen level defined by FiO_2 (as defined and set by trialists)
   c. difference between groups (as defined by trialists)
3. 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%
4. 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%

5. 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
6. According to oxygen delivery system:
   a. invasive mechanical ventilation with endotracheal tube
   b. any non-invasive oxygen administration
   c. mixed oxygen delivery system

Sensitivity analysis

To assess the potential impact of bias, we planned to conduct a sensitivity analysis for each outcome including only trials at overall 'low 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.

Due to a low number of trials reporting on proportion of patients with one or more SAE as previously defined, we conducted two post-hoc defined sensitivity analyses of the reporting of serious adverse events:

1. Estimating the proportion of participants with one or more SAE as the highest reported proportion of specific serious adverse event reported in each trial divided by the number of participants in each intervention group.
2. Estimating the proportion of participants with one or more SAE as the cumulated number of serious adverse events reported in each trial divided by the number of participants in each intervention group.

Summary of findings and assessment of the certainty of the evidence

We used the GRADEpro GDT system (GRADEpro GDT) 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) and constructed summary of findings tables (Guyatt 2008); one including data only from trials at overall low risk of bias and one including data from 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-trial risk of bias, directness of the evidence, heterogeneity of the data, precision of effect estimates (Jakobsen 2014a), and risk of publication bias.

RESULTS

Description of studies

The identified studies are described below.

Results of the search

The searches have been run several times and records imported to Covidence during the process of the first version and this update the review. In total 115.857 have been imported.

For the first version of this review a total of 32,813 titles and abstracts were screened, which entailed forward and backward citation searches, clinical trials registers, and grey literature. Of these, a total of 303 full-text records were assessed, excluding 293, resulting in 10 trials included in the qualitative synthesis and 7 trials in the quantitative synthesis.

In this updated review, a total of 13,509 new titles and abstracts were screened, which entailed forward and backward citation searches, clinical trials registers, and grey literature.

In total, we obtained 567 full-text reports to assess eligibility and excluded 560 references (146 wrong intervention, 153 wrong patient population (34 key studies), 128 duplicate full-

text, 64 wrong study design, 59 wrong publication type, 10 ongoing trials (ACTRN12620000391976; ChiCTR-INR-17012800; ChiCTR-IOR-17011717; CTRI/2020/12/029614; ISRCTN13384956; NCT02999932; NCT03141099; NCT04198077; NCT04425031; NCT04824703)) from the meta-analyses. In total, 567 full-text reports were assessed for eligibility. From these 560 were excluded from the meta-analyses (146 wrong intervention, 153 wrong patient population (34 key studies), 128 duplicate full-text, 64 wrong study design, 59 wrong publication type, 10 ongoing trials (ACTRN12620000391976; ChiCTR-INR-17012800; ChiCTR-IOR-17011717; CTRI/2020/12/029614; ISRCTN13384956; NCT02999932; NCT03141099; NCT04198077; NCT04425031; NCT04824703)).

Of the 10 RCTs identified in the original review we excluded one report (Young 2017) due to overlap in patient population with a new report (Mackle 2020). In all, we included 16 RCTs involving a total of 6486 participants randomly assigned to a higher versus lower oxygenation strategies in the qualitative synthesis (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Ishii 2018; Jakkula 2018; Jun 2019; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Taher 2016; Yang 2019), and 14 reports in the quantitative synthesis (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Jun 2019; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019). Detailed descriptions of included trials are shown in the Characteristics of included studies table.

One trial was identified after the systematic literature search (Gelissen 2021).

For the study flow diagram see Figure 1.

Figure 1.

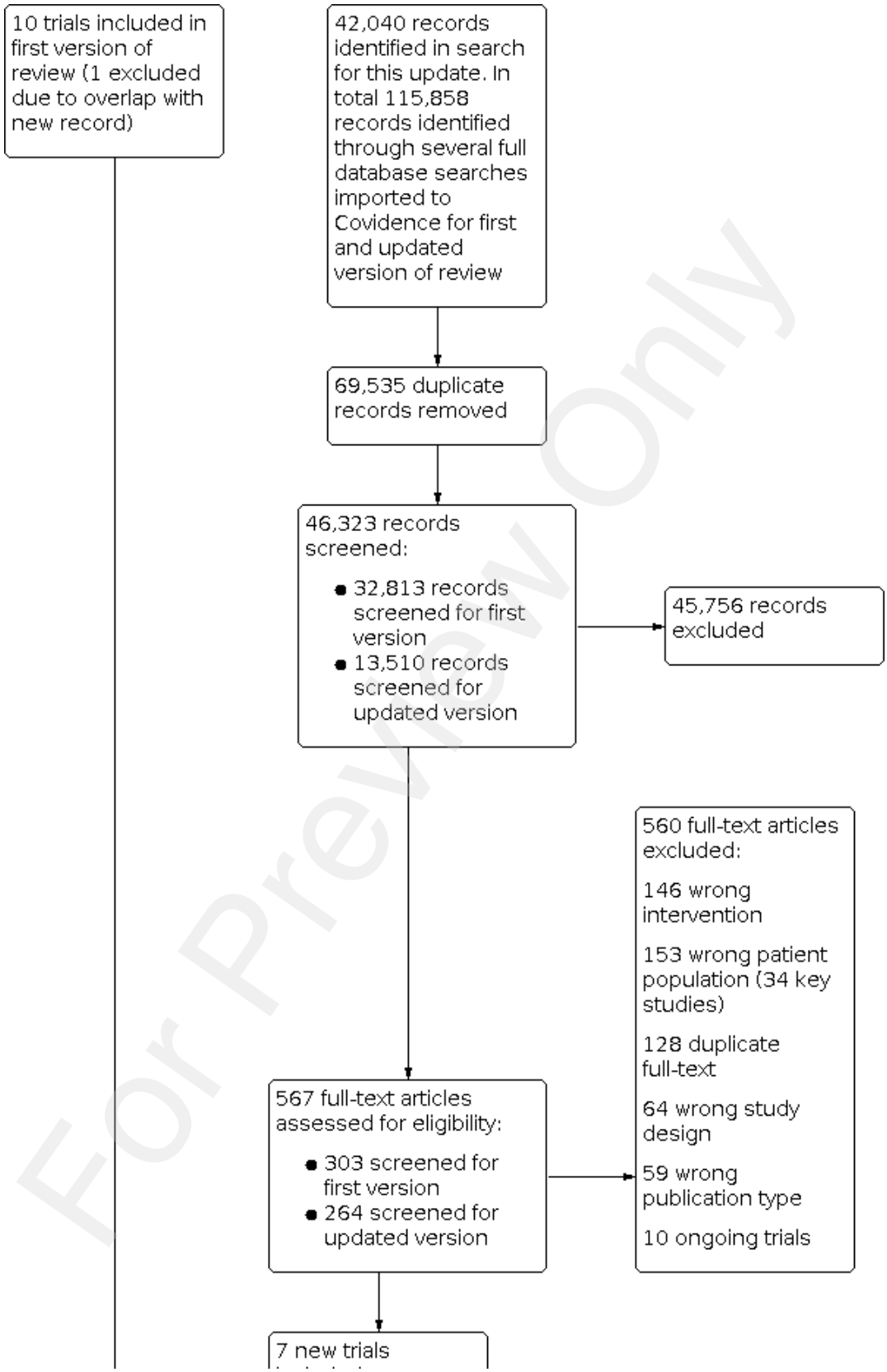
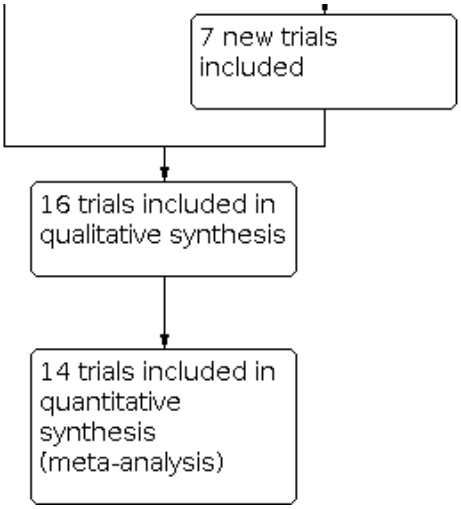


Figure 1. (Continued)



We approached the corresponding authors to request missing or unclear information and received a reply from six authors.

Included studies

Trial characteristics

A total of 6486 participants were randomized in the 16 included trials. See Characteristics of included studies.

Reporting of outcomes

Thirteen trials reported on mortality (6362 participants) (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019).

Three trials reported on the proportion of participants with one or more serious adverse events (SAE) or any SAE (3944 participants) (Asfar 2017; Gelissen 2021; Schjørring 2021). Fourteen trials reported on individual SAEs (6449 participants) (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Jun 2019; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019). Of additional SAEs, two trials reported on delirium (239 participants) (Barrot 2020; Martin 2021), two trials reported on pneumothorax (647 participants) (Asfar 2017; Barrot 2020), three trials reported on intestinal ischemia (3575 participants) (Asfar 2017; Barrot 2020; Schjørring 2021), two trials reported on cardiovascular failure including shock (3408 participants) (Girardis 2016; Schjørring 2021), two trials reported on cardiac arrhythmia (239 participants) (Barrot 2020; Martin 2021), two trials reported on liver failure (1054 participants) (Gelissen 2021; Girardis 2016), and three trials reported on renal failure (1088 participants) (Gelissen 2021; Girardis 2016; Martin 2021). Digestive haemorrhage, digital ischaemia, respiratory failure, seizure, severe hypercapnia and respiratory acidosis, unexplained brain oedema, and ventricular arrhythmias were only reported in single trials (Table 1).

No trials reported on the proportion of participants with lung injury as a composite outcome. Four trials reported on ARDS (871 participants) (Gelissen 2021; Jakkula 2018; Lång 2018; Panwar 2016), and four trials reported on pneumonia (1197 participants) (Asfar 2017; Barrot 2020; Girardis 2016; Lång 2018). No trials reported on pulmonary fibrosis.

Three trials reported on myocardial infarction (3368 participants) (Gelissen 2021; Jun 2019; Schjørring 2021).

Four trials reported on stroke (4476 participants) (Barrot 2020; Gelissen 2021; Mackle 2020; Schjørring 2021).

Two trials reported on sepsis (685 participants) (Barrot 2020; Girardis 2016).

Two trials did not report on any of our pre-defined outcomes or on any serious adverse events (Ishii 2018; Taher 2016).

Trial design

Thirteen trials used a two-arm, parallel-group design (Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Ishii 2018; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Taher 2016; Yang 2019), one trial used a two-factorial design (Asfar 2017), one trial used a two times three factorial design (Jakkula 2018), and one trial used a three-arm design (Jun 2019). The trials were published from 2002 to 2021. Eight trials were conducted in Europe (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Jakkula 2018; Lång 2018; Martin 2021; Schjørring 2021); two in China (Jun 2019; Yang 2019); two in Iran (Mazdeh 2015; Taher 2016); one in Australia and New Zealand (Mackle 2020); one in Australia, New Zealand, and France (Panwar 2016); one in Hong Kong (Gomersall 2002); and one in Japan (Ishii 2018).

Three trials were classified as either feasibility (Martin 2021) or pilot trials (Lång 2018; Panwar 2016).

Loss to follow-up

Loss to follow-up varied among trials, the lowest being 0% in three trials (Martin 2021; Taher 2016; Yang 2019), and the highest being just over 30% (Gelissen 2021).

Of the 16 included trials, 11 trials had less than 5% loss to follow-up (Asfar 2017; Barrot 2020; Girardis 2016; Jakkula 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Taher 2016; Yang 2019), two trials had 5 to 10% loss to follow-up (Gomersall 2002; Lång 2018), one trial had 14% loss to follow-up (Ishii 2018), and one trial more than 30% loss to follow-up (Gelissen 2021).

Loss to follow-up could not be ascertained for one trial due to limited information reported (Jun 2019).

Participants

Number of participants

The number of participants in the trials ranged from 34 to 2928. The approximate mean age of participants was 61 years, and the approximate mean proportion of male participants was 65%.

Types of intensive care units

All trials included adults admitted to the ICU: nine trials included multidisciplinary ICU-patients (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019); three trials included only medical ICU-patients (Gomersall 2002; Jakkula 2018; Mazdeh 2015); and two trials included only surgical ICU-patients (Ishii 2018; Lång 2018). Two trials included only adults with traumatic brain injury (Lång 2018; Taher 2016); one trial only adults after resuscitation from out-of-hospital cardiac arrest (Jakkula 2018); and one trial only adults with stroke (Mazdeh 2015). Two trials failed to report the type of ICU to which patients were admitted (Jun 2019; Taher 2016).

Oxygen delivery systems

Eight trials included only adults receiving invasive mechanical ventilation (Asfar 2017; Barrot 2020; Ishii 2018; Jakkula 2018; Lång 2018; Martin 2021; Panwar 2016; Taher 2016); two trials included only adults receiving any non-invasive oxygen administration (Gomersall 2002; Mazdeh 2015); and six trials included adults on both invasive mechanical ventilation and adults receiving non-invasive oxygen administration (Gelissen 2021; Girardis 2016; Jun 2019; Mackle 2020; Schjørring 2021; Yang 2019).

Respiratory failure

Three trials restricted inclusion to patients with a PaO_2/FiO_2 ratio ≥ 100 mmHg (Asfar 2017; Gelissen 2021; Jakkula 2018), one trial restricted inclusion to patients with a PaO_2/FiO_2 ratio ≥ 150 mmHg (Girardis 2016), and one trial restricted inclusion to patients not being hypoxic or for whom oxygen therapy was inevitable (selection criteria not further specified by trialists) (Mazdeh 2015). One trial excluded patients with a PaO_2 < 13 kPa or an SpO_2 < 95% with an FiO_2 of 0.40 and positive end-expiratory pressure (PEEP) of 10 cm H_2O or oxygenation failure was judged probable during ICU admission (Lång 2018).

Three trials included only patients with respiratory failure: one trial restricted inclusion to patients fulfilling the ARDS criteria (Barrot 2020; ARDS Definition Task Force 2012); one trial required patients to receive ≥10 litres of oxygen per minute in an open system or an FiO_2 ≥ 0.50 in a closed system (Schjørring 2021); and one trial required the diagnosis of respiratory failure (as defined by clinicians) (Martin 2021).

Four trials excluded patients with either known chronic obstructive pulmonary disease (COPD) (Barrot 2020; Jakkula 2018) or acute decompensation of COPD (Girardis 2016; Yang 2019). Two trials restricted inclusion to patients with COPD (Gomersall 2002; Jun 2019). One trial excluded patients with known severe COPD (Gelissen 2021). One trial excluded patients with known (or being highly suspected to having) chronic lung disease with a baseline SpO_2 in the range of 88-92% (Martin 2021).

Hypoxaemic encephalopathy or cerebral pathology

Five trials excluded participants resuscitated from cardiac arrest prior to randomization (Asfar 2017; Barrot 2020; Martin 2021; Mazdeh 2015; Taher 2016), whilst one trial restricted inclusion to patients resuscitated from witnessed out of hospital cardiac arrest (Jakkula 2018).

Four trials excluded participants with intra-cranial pathology prior to randomization being either: intra-cranial hypertension (Asfar 2017); intra-cranial hypertension or traumatic brain injury (Barrot 2020); confirmed or suspected acute or pre-existing intra-cranial pathology or suspicion of increased intra-cranial pressure, or both (Jakkula 2018); or any penetrating traumatic brain injury (Lång 2018). Three trials restricted inclusion to patients with cerebral pathology only (Lång 2018; Mazdeh 2015; Taher 2016).

Three trials included participants with any risk factors for either hypoxaemic encephalopathy (e.g. cardiac arrest prior to randomization) or any cerebral pathology (e.g. traumatic brain injury) (Gelissen 2021; Mackle 2020; Schjørring 2021).

Limitations of care

Ten trials excluded participants with limitations of care or with short remaining life-expectancy as evaluated by clinicians (Asfar 2017; Barrot 2020; Girardis 2016; Gomersall 2002; Lång 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019).

Co-enrolment

Eight trials explicitly co-enrolled participants into other clinical trials (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Mackle 2020; Martin 2021; Schjørring 2021; Yang 2019).

Pregnancy

Twelve trials explicitly excluded pregnant participants (Asfar 2017; Barrot 2020; Gelissen 2021;Girardis 2016; Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Taher 2016; Yang 2019).

Haemodynamic insufficiency

One trial restricted inclusion to participants with refractory septic shock (Asfar 2017), whilst one trial required participants to be haemodynamically stabile (Taher 2016), and one trial excluded participants with either shock or malignant arrhythmia (Jun 2019).

Time limits to inclusion

Nine trials had a minimum expectation to patients' length of stay in the ICU: one trial expected mechanical ventilation ≥12 hours (Ishii

2018); one trial expected ≥24 hours of oxygen supplementation in the ICU (Schjørring 2021); one trial expected mechanical ventilation beyond the next calendar day (Mackle 2020); two trials expected mechanical ventilation ≥24 hours (Lång 2018; Panwar 2016); one trial expected ICU-stay ≥48 hours (Gelissen 2021); and three trials expected ≥72 hours of mechanical ventilation (Girardis 2016; Martin 2021; Yang 2019).

One trial excluded participants with more than two hours of invasive mechanical ventilation or non-invasive mechanical ventilation or both, in an ICU during current hospital admission (Mackle 2020); one trial restricted inclusion to less than or equal to six hours from start of vasopressors (Asfar 2017); four trials had a time limit of 12 hours to inclusion from either ICU-admission (Gelissen 2021; Schjørring 2021; Yang 2019), or to start of invasive mechanical ventilation (Barrot 2020); and one trial required participants to be randomized within 18 hours from ICU-admission and within 36 hours from injury (Lång 2018).

Baseline severity of illness

Scores for baseline disease severity were reported in various manners: four trials reported APACHE II, approximate mean 21.4 (range 17.0 to 23.5) (Lång 2018; Mackle 2020; Martin 2021; Yang 2019); two trials reported SAPS II, approximate mean 33 (range 28 to 38) (Girardis 2016; Jakkula 2018); two trials reported SAPS III, approximate mean 69 (range 67 to 71) (Asfar 2017; Barrot 2020); two trials reported SOFA scores, approximate mean 9 (range 8 to 9) (Barrot 2020; Schjørring 2021); one trial reported SOFA scores excluding the respiratory component, approximate mean 5.5 (range 5 to 6) (Gelissen 2021); one trial reported APACHE III, approximate mean 75 (range 70 to 80) (Panwar 2016); one trial reported Bartel index, mean 42 (Mazdeh 2015); one trial reported Glasgow Coma Scale, approximate mean 7.4 (Taher 2016). Three trials failed to report any illness severity scores (Gomersall 2002; Ishii 2018; Jun 2019).

Lengths of interventions

The maximum duration of the applied intervention varied greatly among trials. The shortest duration was a maximum of 6 hours after randomization (Taher 2016), whilst the longest was a maximum of 90 days after randomization (Schjørring 2021). Six trials failed to report the maximum duration of intervention (Girardis 2016; Ishii 2018; Jun 2019; Martin 2021; Panwar 2016; Yang 2019). Details on interventions are provided in Table 2.

Funding

Twelve trials were funded by public grants (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Taher 2016; Yang 2019); two trials did not report how they were funded (Ishii 2018; Jun 2019); two trials were funded by public and private grants and specified that funding bodies had no input regarding the design, management, or reporting of the trial (Jakkula 2018; Schjørring 2021).

Experimental intervention

Of the 16 included trials, four trials randomized participants to higher versus lower oxygen by using FiO_2 (Jun 2019; Lång 2018; Mazdeh 2015; Taher 2016); ten trials randomized participants to an oxygenation target (or target range) (Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019); and two trials randomized participants to a specific FiO_2 in the experimental group and to target specific oxygenation levels in the control group (Asfar 2017; Ishii 2018).

Of the eight trials using FiO_2 in the experimental group, two trials used an FiO_2 of 1.0 (Asfar 2017; Ishii 2018); one trial used an FiO_2 of 0.80 (Taher 2016); one trial used an FiO_2 of 0.70 (Lång 2018); one trial used an FiO_2 0.40 to 0.70 (Jun 2019); one trial used FiO_2 of 0.50 (Mazdeh 2015); one trial used $FiO_2 \geq 0.40$ (Girardis 2016); and one trial used $FiO_2 \geq 0.30$ (Yang 2019).

In the experimental (higher) group, two trials targeted an SpO_2 of ≥ 96% (Martin 2021; Panwar 2016) one trial targeted a PaO_2 of 12 to 14 kPa (90 to 105 mmHg) or an SpO_2 of ≥96% (Barrot 2020); one trial targeted a PaO_2 of 14 to 18 kPa (105 to 135 mmHg) (Gelissen 2021); one trial targeted an SpO_2 of 97% to 100% (Girardis 2016); one trial targeted a PaO_2 above 9.0 kPa (67.5 mmHg) (Gomersall 2002); one trial targeted a PaO_2 of 20 to 25 kPa (150 to 187.5 mmHg) (Jakkula 2018); 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) (Mackle 2020); and one trial targeted a PaO_2 of 12 kPa (90 mmHg) (Schjørring 2021).

Two trials were categorised by us as using a low target in the experimental (higher) group (Gomersall 2002; Mazdeh 2015), and thirteen trials were categorised as using a high target in the experimental group (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Ishii 2018; Jakkula 2018; Jun 2019; Lång 2018; Martin 2021; Panwar 2016; Schjørring 2021; Taher 2016; Yang 2019). One trial could not be categorised according to our definitions, as no specific target was used (Mackle 2020).

Details on interventions are provided in Table 2.

Comparator intervention

Four trials used an FiO_2 in the control group; one trial used the expected FiO_2 to achieve a PaO_2 of 13.3 kPa (100 mmHg) (Ishii 2018); one trial used an FiO_2 of 0.40 (Lång 2018); one trial used an FiO_2 of 0.30 to 0.50 (Jun 2019); and one trial used an FiO_2 of 0.50 (Taher 2016).

In the control group (lower) three trials targeted an SpO_2 88% to 92% by itself (Martin 2021; Panwar 2016) or in combination with a PaO_2 of 7.3 to 9.3 kPa (55 to 70 mmHg) (Barrot 2020); one trial targeted an SaO_2 between 88% and 95% (Asfar 2017); one trial targeted an SpO_2 between 94% and 98% or a PaO_2 of 9.3 to 13. kPa (70 to 100 mmHg) (Girardis 2016) ; one trial targeted an SpO_2 between 95% and 98% or a PaO_2 of 10 to 15 kPa (75 to 112.5 mmHg) (Jakkula 2018); one trial targeted a PaO_2 of 8 to 12 kPa (60

to 90 mmHg) (Gelissen 2021); one trial targeted a PaO_2 of 8 kPa (60 mmHg) (Schjørring 2021); one trial targeted an SpO_2 between 90% and 95% (Yang 2019); one trial targeted a PaO_2 > 6.6 kPa (50 mmHg) (Gomersall 2002); and one trial used an SaO_2/SpO_2 between 91% to 96% (Mackle 2020). One trial used no supplemental oxygen (Mazdeh 2015).

Nine trials were categorised by us as using a low target in the control group (Asfar 2017; Barrot 2020; Gomersall 2002; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019), and six trials were categorised as using a high target in the control (lower) group (Gelissen 2021; Girardis 2016; Ishii 2018; Jakkula 2018; Jun 2019 Lång 2018; Taher 2016).

Details on interventions are provided in Table 2.

Primary outcomes as defined by trialists

Six trials reported mortality as their primary outcome, at various lengths of follow-up: one within ICU-admission (Girardis 2016), one at 14 days (Jun 2019); three at 28 days (Asfar 2017; Barrot 2020; Yang 2019), and one at 90 days post-randomisation (Schjørring 2021).

Two trials reported the Bartel index at hospital discharge and 6 months (Taher 2016), and also at the first day of admission (Mazdeh 2015), as the primary outcome.

Two trials reported feasibility measures (as defined by trialists) as the primary outcome (Martin 2021; Panwar 2016).

One trial reported the number of ventilator-free days at day 28 post-randomization as the primary outcome (Mackle 2020).

One trial reported the occurrence af atelectases within five days of randomization as the primary outcome (Ishii 2018).

One trial reported the need for mechanical ventilation or death within current hospital admission as the primary outcome (Gomersall 2002).

One trial reported the serum concentration of neuro-specific enolase (NSE) 48 hours after cardiac arrest as the primary outcome (Jakkula 2018).

One trial reported a collection of laboratory markers (level of reactive oxygen species, inter-leukin-6 and NSE) during ICU-care as the primary outcome (Lång 2018).

One trial reported the cumulative daily delta Sequential Organ Failure Assessment (SOFA) score (omitting sub-scores from the respiratory component) from day 1 to 14 ($SOFA_{rank}$) (Gelissen 2021).

Mortality as an outcome

Mortality was also reported as a secondary outcome at various lengths of follow-up: five trials reported on ICU-mortality (Barrot 2020; Gelissen 2021; Martin 2021; Panwar 2016; Yang 2019), four trials reported on hospital mortality (Gelissen 2021; Girardis 2016; Gomersall 2002; Panwar 2016), three trials reported on 90-day mortality (Gelissen 2021; Mackle 2020; Panwar 2016), and two trials reported on 180-day mortality (Lång 2018; Mackle 2020).

Excluded studies

We excluded RCTs of higher versus lower oxygenation strategies that were conducted in populations not being admitted to an ICU. We listed the reasons for exclusion of 35 key excluded trials, which included RCTs of higher versus lower oxygenation strategies for participants who were acutely ill but not admitted to the ICU, as detailed in the Characteristics of excluded studies table.

Ongoing studies

We identified ten ongoing trials (ACTRN12620000391976; ChiCTR-INR-17012800; ChiCTR-IOR-17011717; CTRI/2020/12/029614; ISRCTN13384956; NCT02999932; NCT03141099; NCT04198077; NCT04425031; NCT04824703) , which may be included in future updates of this review. See Characteristics of ongoing studies.

Risk of bias in included studies

We assessed all trials according to the Risk of Bias 2 tool (Higgins 2016; Sterne 2019) for all reported outcomes.

Two trials were not evaluated in terms of 'Risk of Bias', as none of the predefined outcomes were reported on (Ishii 2018; Taher 2016).

Risk of bias tables are presented along with all meta-analyses.

A) Domain 1: Randomization process

Eleven trials described the generation of the allocation sequence adequately, using computer-generated random numbers, and were judged to be at “low risk of bias” for this domain (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). One trial stated that the trial was randomized and used sealed, opaque envelopes, but the method of sequence generation was not described (Lång 2018). We judged this trial to be at “low risk of bias'. Two trials stated that the trial was randomized, but the method of sequence generation was not described resulting in “some concerns of bias” (Jun 2019; Mazdeh 2015).

B) Domain 2: Deviations from the intended interventions

Eight trials were judged to be at “low risk of bias” for this domain (Asfar 2017; Barrot 2020; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). Two trials were judged to be at “some concerns of bias” for this domain; one due to the use of an unjustified, modified intention-to-treat analysis (Girardis 2016); and one due to no information on blinding, protocol deviations, group allocation numbers, and missing event rates (Jun 2019). Four trials were judged to be at “high risk of bias” for this domain; one trial excluded 1 patient post-randomization due to development of severe ARDS (Gelissen 2021); one trial due to use of modified intention-to-treat analyses excluding two patients post-randomization who violated inclusion criteria (Gomersall 2002); one trial did not specify group allocation of five patients lost to follow-up (Lång 2018); and one trial did not report on the analyses used to estimate the effect of the intervention (Mazdeh 2015).

C) Domain 3: Missing outcome data

Ten trials were judged to be at "low risk of bias" for this domain (Asfar 2017; Barrot 2020; Girardis 2016; Jakkula 2018; Jun 2019; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). Four trials were judged to be at "high risk of bias" for this domain; one trial had more than 30% loss to follow-up (Gelissen 2021); one trial had more than 5% lost to follow-up and no description on handling of missing data (Gomersall 2002); one trial had five patients lost to follow-up, but did not report on reasons for loss to follow-up nor group allocation (Lång 2018); and one trial had one patient lost to follow up that was not described in the report, and did not report on the patient's group allocation (Mazdeh 2015).

D) Domain 4: Measurement of the outcome

All fourteen trials reporting data for this review were judged to be at "low risk of bias" for this domain.

E) Domain 5. Selection of the reported results

Nine trials were judged to be at "low risk of bias" for this domain (Asfar 2017; Barrot 2020; Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). Four trials were judged to be at "some concerns of risk of bias" for this domain; we were unable to find a trial protocol for two trials (Gomersall 2002; Jun 2019); one trial reported 90-day mortality post-hoc, and did not pre-specify the reporting of SAEs (Gelissen 2021); and one trial was retrospectively registered (Mazdeh 2015). One trial was judged to be at "high risk of bias" for this domain as the trial reported the results from an unjustified, modified intention-to-treat analysis as the primary outcome, and the trial was judged being registered retrospectively (Girardis 2016).

F) Overall risk of bias

Mortality

Thirteen trials reported on mortality, of which eight were judged to be at overall low risk of bias in all domains for this outcome (Asfar 2017; Barrot 2020; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). The remaining five trials were judged to be at overall high risk of bias (Gelissen 2021; Gomersall 2002; Lång 2018; Mazdeh 2015).

Serious adverse events

Fourteen trials reported on the proportion of participants with one or more SAE or any single SAE, of which eight were judged to be at overall low risk of bias in all domains for this outcome (Asfar 2017; Barrot 2020; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). The remaining six trials were judged to be at overall high risk of bias (Gelissen 2021; Girardis 2016; Gomersall 2002; Jun 2019; Lång 2018; Mazdeh 2015).

Quality of life

One trial reported on quality of life (Mackle 2020). For this outcome, the trial was judged to be at overall high risk of bias.

Lung injury

Seven trials reported on any lung injury (defined as either ARDS, pneumonia, or pulmonal fibrosis), of which three trials were judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020; Jakkula 2018; Panwar 2016). The remaining four trials were judged to be at overall high risk of bias for this outcome (Asfar 2017; Gelissen 2021; Girardis 2016; Lång 2018).

Three trials reported on ARDS, of which two trials were judged to be at overall low risk of bias in all domains for this outcome (Jakkula 2018; Panwar 2016). The other trials was judged to be at overall high risk of bias for this outcome (Lång 2018).

Four trials reported on pneumonia, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020). One trial was judged to be at some concerns for this outcome (Asfar 2017), and two trials were judged to be at overall high risk of bias for this outcome (Girardis 2016; Lång 2018).

Myocardial infarction

Three trials reported on myocardial infarction, of which 1 trial was judged to be at overall low risk of bias in all domains for this outcome (Schjørring 2021). The remaining two trials were judged to be at overall high risk of bias for this outcome (Gelissen 2021; Jun 2019).

Stroke

Four trials reported on stroke, of which two trials were judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020; Schjørring 2021), one trial was judged to be at some concerns for this outcome (Mackle 2020), and one trial was judged to be at overall high risk of bias for this outcome (Gelissen 2021).

Sepsis

Two trials reported on sepsis, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020). The other trials was judged to be at overall high risk of bias for this outcome (Girardis 2016).

Additional serious adverse events

Two trials reported on delirium, and both trials were judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020; Martin 2021).

Two trials reported on pneumothorax, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020). The other trial was judged to be at some concerns for this outcome (Asfar 2017).

Three trials reported on intestinal ischaemia, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Schjørring 2021). The other trials were judged to be at some concerns for this outcome (Asfar 2017; Barrot 2020).

Two trials reported on cardiovascular failure including shock, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Schjørring 2021). The other trial was judged to be at overall high risk of bias for this outcome (Girardis 2016).

Two trials reported on cardiac arrhythmia, and both trials were judged to be at overall low risk of bias in all domains for this outcome (Barrot 2020; Martin 2021).

Two trials reported on liver failure, and both trials were judged to be at overall high risk of bias for this outcome (Gelissen 2021; Girardis 2016).

Three trials reported on renal failure, of which one trial was judged to be at overall low risk of bias in all domains for this outcome (Martin 2021). The other two trials were judged to be at overall high risk of bias (Gelissen 2021; Girardis 2016).

Effects of interventions

See: **Summary of findings 1** Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU — trials at overall low risk of bias only; **Summary of findings 2** Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the ICU — all included trials

Primary outcomes

All-cause mortality — at maximum follow-up

Thirteen of the 16 trials included reported on all-cause mortality, with a total of 6262 participants randomized and a mean follow-up of 4 months (range 1 to 6 months) (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019).

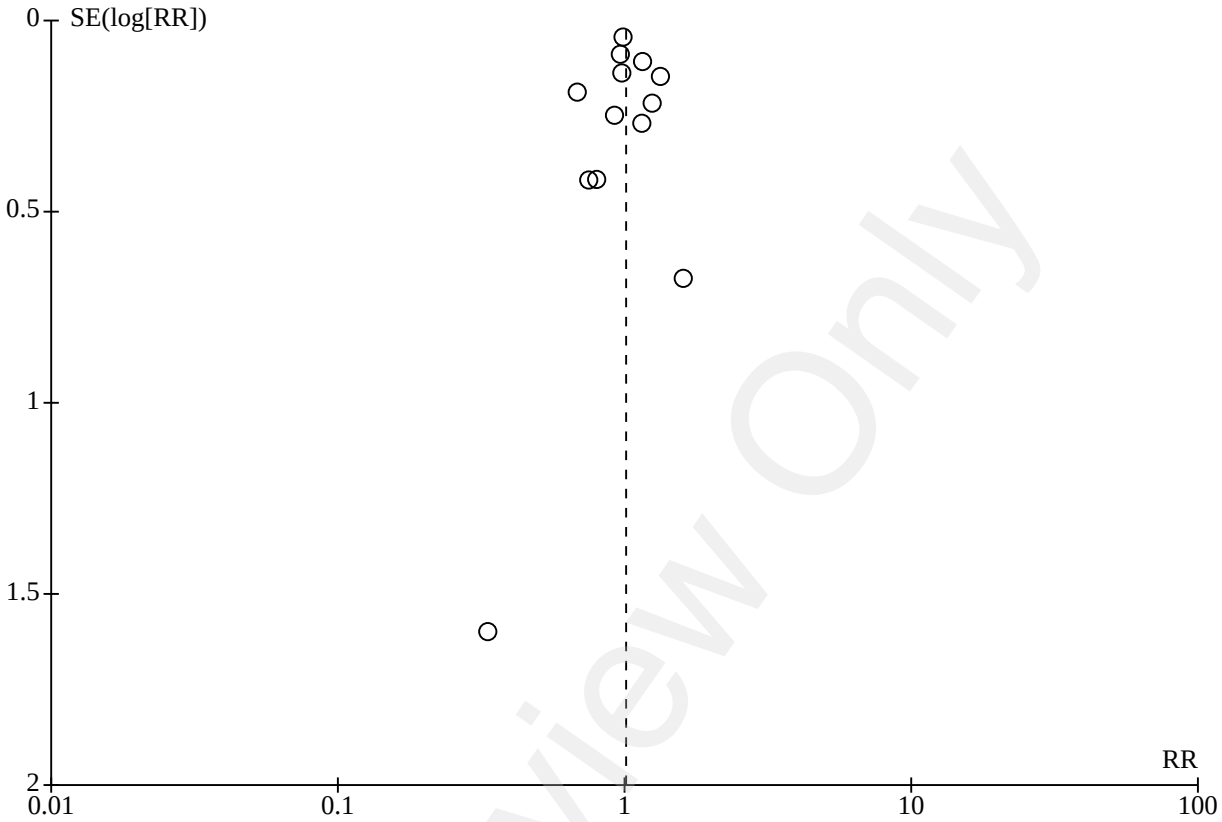
Eight trials were judged to be at overall low risk of bias (5050 participants) (Asfar 2017; Barrot 2020; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). A total of 40.0% in the higher group versus 40.4% in the lower group had died. In these trials, meta-analysis of all-cause mortality indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (random-effects model RR 0.99, 95% CI 0.91 to 1.09; I^2 = 13%; 4945 participants; 8 trials; Analysis 1.1; low certainty evidence; Summary of findings 1).

In all trials, a total of 38.5% in the higher group versus 38.3% in the lower group died. Meta-analysis of all-cause mortality in all trials did not show any difference in effect of higher versus lower oxygenation strategies (random-effects model risk ratio (RR) 1.01, 95% confidence interval (CI) 0.94 to 1.10; I^2 = 9%; 5973 participants; 13 trials; Analysis 2.1; very low certainty evidence; Summary of findings 2).

Publication bias

Funnel plot of all-cause mortality at maximum follow-up is presented in Figure 2. No small-study effect was indicated by the Harbord test (P = 0.52) or Begg's test (P = 0.58).

Figure 2. Funnel plot of the risk of mortality at maximum follow-up. A relative risk (RR) <1 indicates benefit of higher oxygenation strategies, whilst an RR >1 indicate benefit of lower. Each circle represents the point estimate of the trials. The black dashed line represent the point estimate the RR of all-cause mortality at maximum follow-up (1.01). Abbreviations: log: natural logarithm; SE: standard error; RR: relative risk.



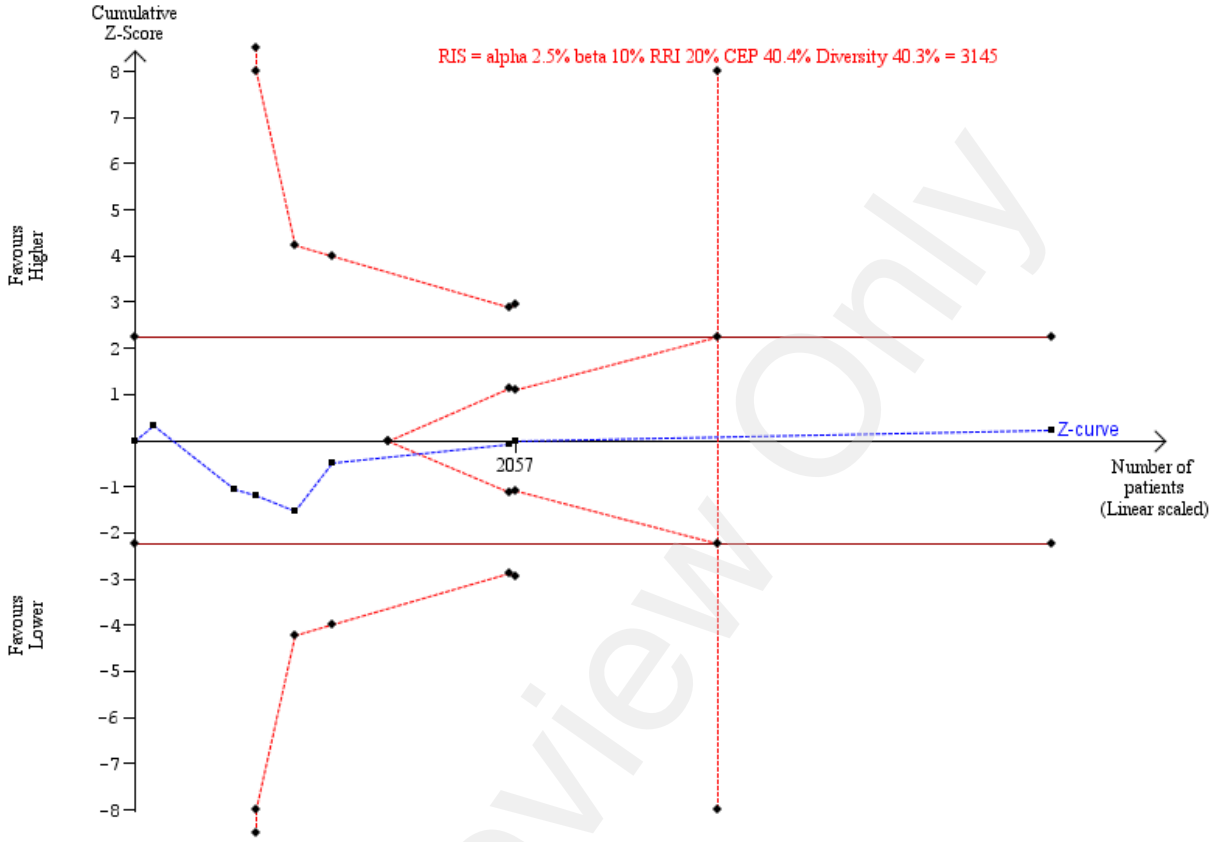
Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 9%; Chi^2 = 13.19, P = 0.36) indicated statistical heterogeneity.

Trial Sequential Analyses

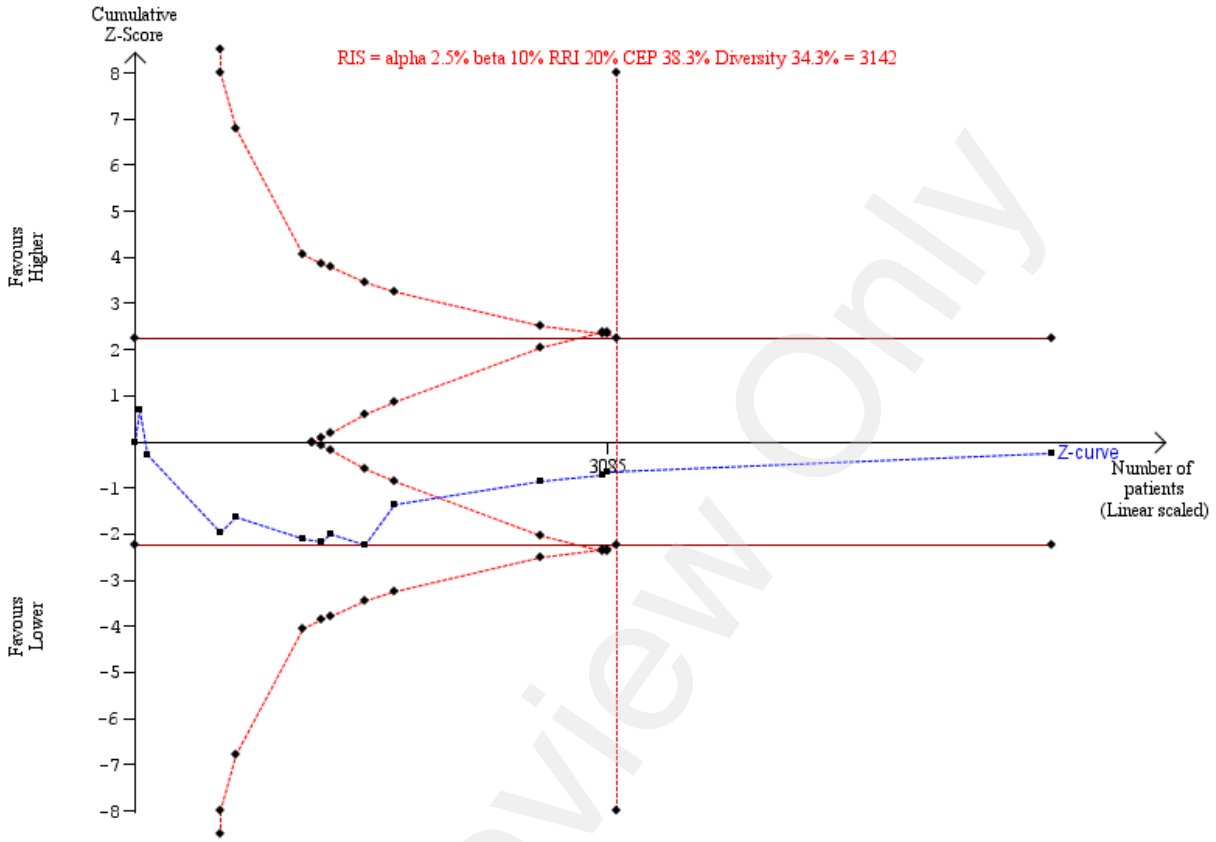
Trial Sequential Analysis of trials judged to be at overall low risk of bias showed that with an anticipated RRI of 20%, mortality in the control group of 40.4%, a type 1 error level of 2.5%, a type two error level of 10%, and a diversity of 40.3%, the required information size was 3,145 participants. The cumulated Z-curve crossed the trial sequential monitoring boundaries for futility, and with 4945 participants in the analysis the required information size was exceeded. This indicated that, considering repetitive testing, the evidence was sufficient to refute a 20% RRI or a 20% RRR for benefit or harm of higher versus lower oxygenation strategies (Figure 3). The TSA CI was 0.90 to 1.10.

Figure 3. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the risk of mortality in trials judged to be at overall low risk of bias. The analysis was based on a mortality in the control group (control event proportion = CEP) of 40.4%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 40.3%. The cumulative Z-curve crossed the trial sequential monitoring boundaries for futility, and the required information size (RIS) was exceeded.



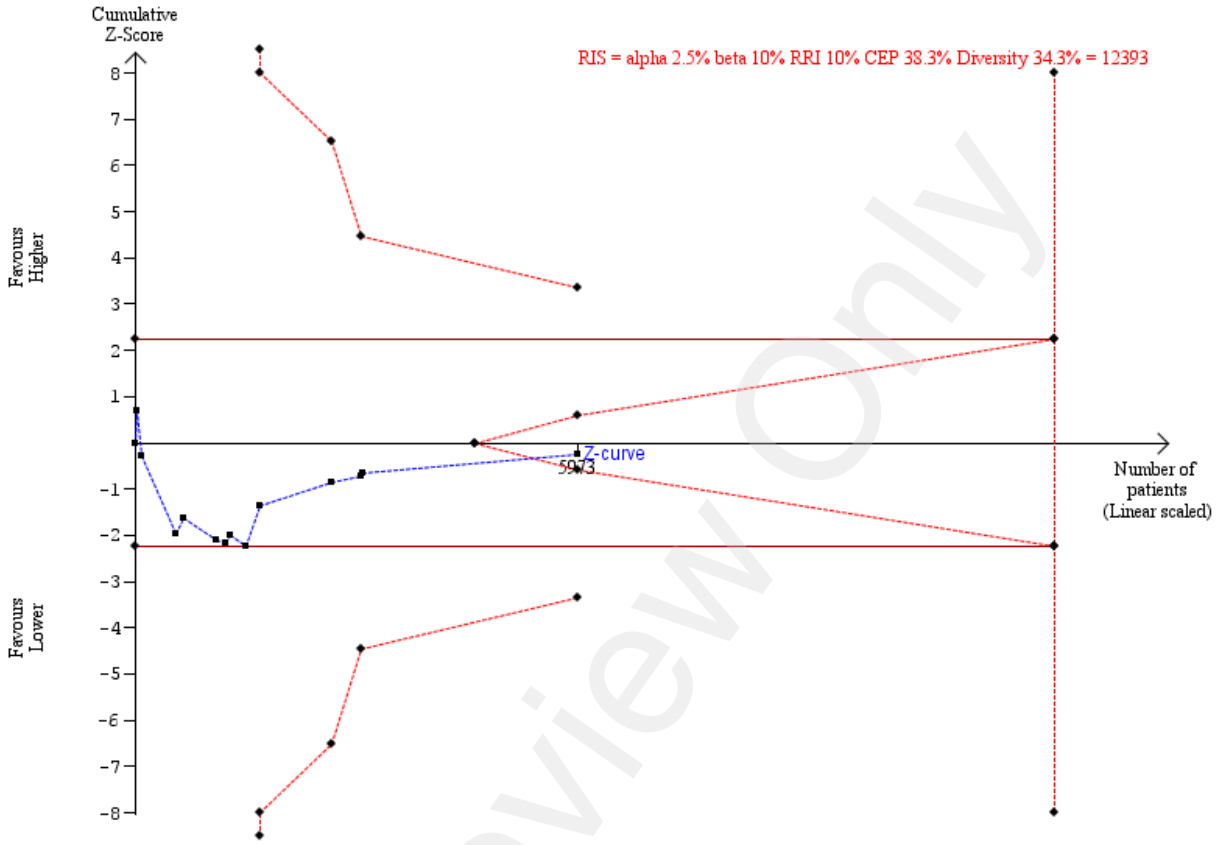
Trial Sequential Analysis of all trials showed that with an anticipated RRI of 20%, mortality in the control group of 38.3%, a type 1 error level of 2.5%, a type two error level of 10%, and a diversity of 34.3%, the required information size was 3,142 participants. The cumulated Z-curve crossed the trial sequential monitoring boundaries for futility, and with 5973 participants in the analysis the required information size was exceeded. This indicated that, considering repetitive testing, the evidence was sufficient to refute a 20% RRI or a 20% RRR for benefit or harm of higher versus lower oxygenation strategies (Figure 4). The TSA CI 0.93 to 1.09.

Figure 4. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the risk of mortality. The analysis was based on a mortality in the control group (control event proportion = CEP) 38.3%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 34.3%. The cumulative Z-curve crossed the trial sequential monitoring boundaries for futility, and the required information size (RIS) was exceeded.



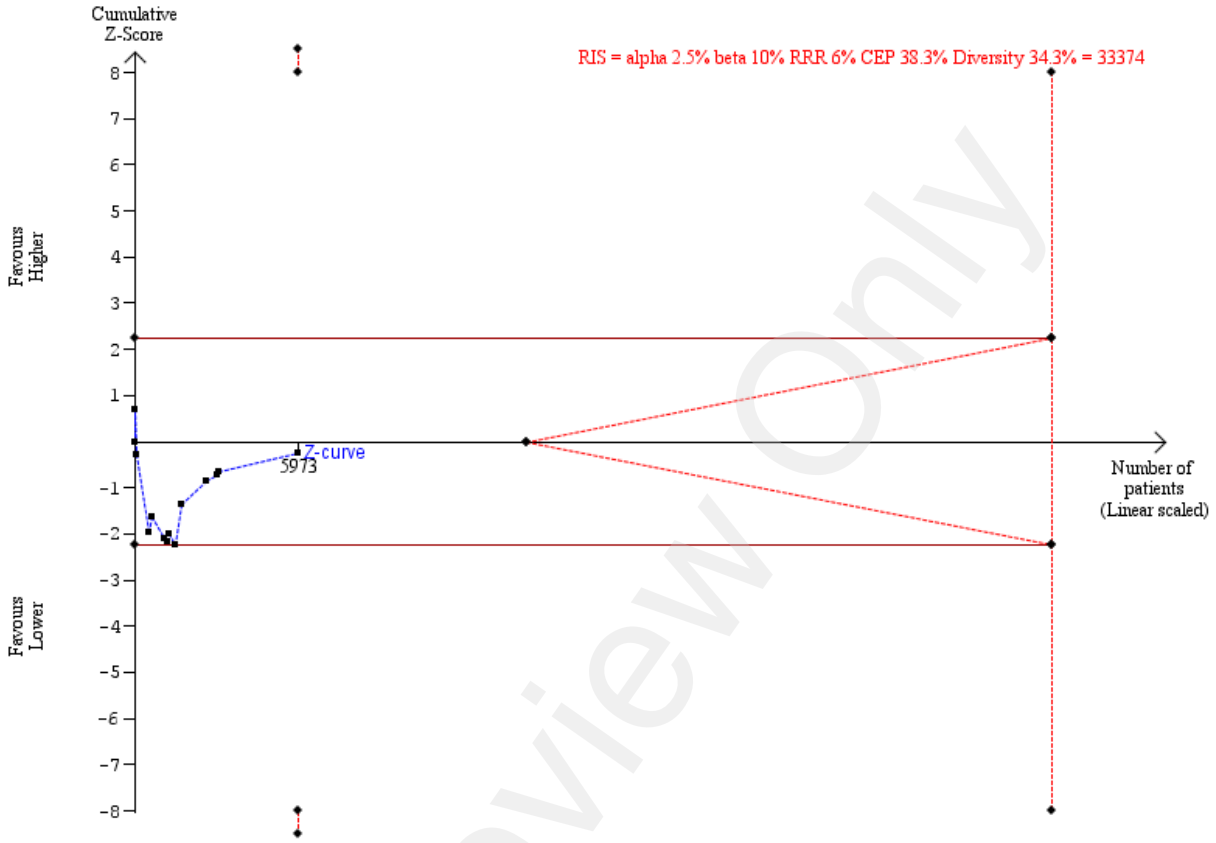
Trial Sequential Analysis with an anticipated RRI of 10%, mortality in the control group of 38.3%, a type 1 error level of 2.5%, a type two error level of 10%, and a diversity of 34.3%, the required information size was 12,393 participants. The cumulative Z-curve crossed trial sequential monitoring boundaries for futility. This indicated that considering sparse data and repetitive testing, evidence was sufficient to refute a 10% RRR or a 10% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 5). The TSA CI was 0.91 to 1.12.

Figure 5. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the risk of mortality. The analysis was based on a mortality in the control group (control event proportion = CEP) of 38.3%, a relative risk increase (RRI) of 10%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 34.3%. Required information size = RIS. The cumulative Z-curve crossed the trial sequential monitoring boundaries for futility.



Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of the analysis of mortality closest to the null-effect (RRR of 6%), mortality in the control group of 38.3%, a type 1 error level of 2.5%, a type two error level of 10%, and a diversity of 34.3%, the required information size was 33,374 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 6% RRR or a 6% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 6). The TSA CI 0.78 to 1.31.

Figure 6. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the risk of mortality. The analysis was based on a mortality in the control group (control event proportion = CEP) of 38.3%, a relative risk reduction (RRR) of 6%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 34.3%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

We found no evidence of a difference in the subgroup analyses according to: risk of bias (Analysis 2.2); different types of oxygen interventions (Analysis 2.3); oxygenation target in the higher oxygen-administration group (Analysis 2.4); FiO_2 or oxygenation target in the lower oxygen-administration group (Analysis 2.5); or oxygen delivery system (Analysis 2.7). When considering effects according to ICU population only subtotals are presented as more than one trial is represented in more than one subgroup (Analysis 2.6).

Sensitivity analyses

The sensitivity analysis excluding trials comparing two low oxygenation strategies or two high oxygenation strategies indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies on all-cause mortality (RR 0.99, 95% CI 0.92 to 1.07; I^2 = 35%; 3874 participants; 6 trials; Analysis 2.8).

The sensitivity analysis assessing the impact of missing data indicated that incomplete outcome did not have the potential to influence the results:

- Best-worst-case scenario random-effects meta-analysis: RR 0.88, 95% CI 0.73 to 1.06; I^2 = 79%; 6261 participants; 13 trials; Analysis 2.9);
- Worst-best-case scenario random-effects meta-analysis: RR 1.13, 95% CI 0.99 to 1.39; I^2 = 74%; 6261 participants; 13 trials; Analysis 2.10).

Data were imputed for eleven trials (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Jakkula 2018; Lång 2018; Mackle 2020; Mazdeh 2015; Panwar 2016; Schjørring 2021).

Proportion of participants with one or more SAE

Three of 16 trials (3944 participants) reported on the proportion of participants with one or more SAE as a composite outcome as according to our definition (Asfar 2017; Gelissen 2021; Schjørring 2021). Two of these trials were judged to be at overall low risk of bias (3370 participants) (Asfar 2017; Schjørring 2021).

In trials judged to be at overall low risk of bias, a total of 44.2% in the higher group versus 41.3% in the lower group had at least one SAE. In these trials, meta-analysis of proportion of patients with one or more SAE indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (fixed-effect model RR 1.07, 95% CI 0.99 to 1.15; I^2 = 6%; 3344; 2 trials; Analysis 1.2; low certainty evidence; Summary of findings 1).

In all trials, a total of 43.8% in the higher group versus 41.5% in the lower group had at least one SAE. Meta-analyses of the proportion of participants with one or more SAEs in all trials showed no difference in the effect of higher versus lower oxygenation strategies (fixed-effect model RR 1.05, 95% CI 0.98 to 1.13; I^2 = 27%; 3744 participants; 3 trials; Analysis 3.1; low certainty evidence; Summary of findings 2).

Publication bias

No evaluation of publication bias was performed the pre-defined outcome of the proportion of participants with one or more SAE as less than 10 trials reported on this outcome.

Heterogeneity

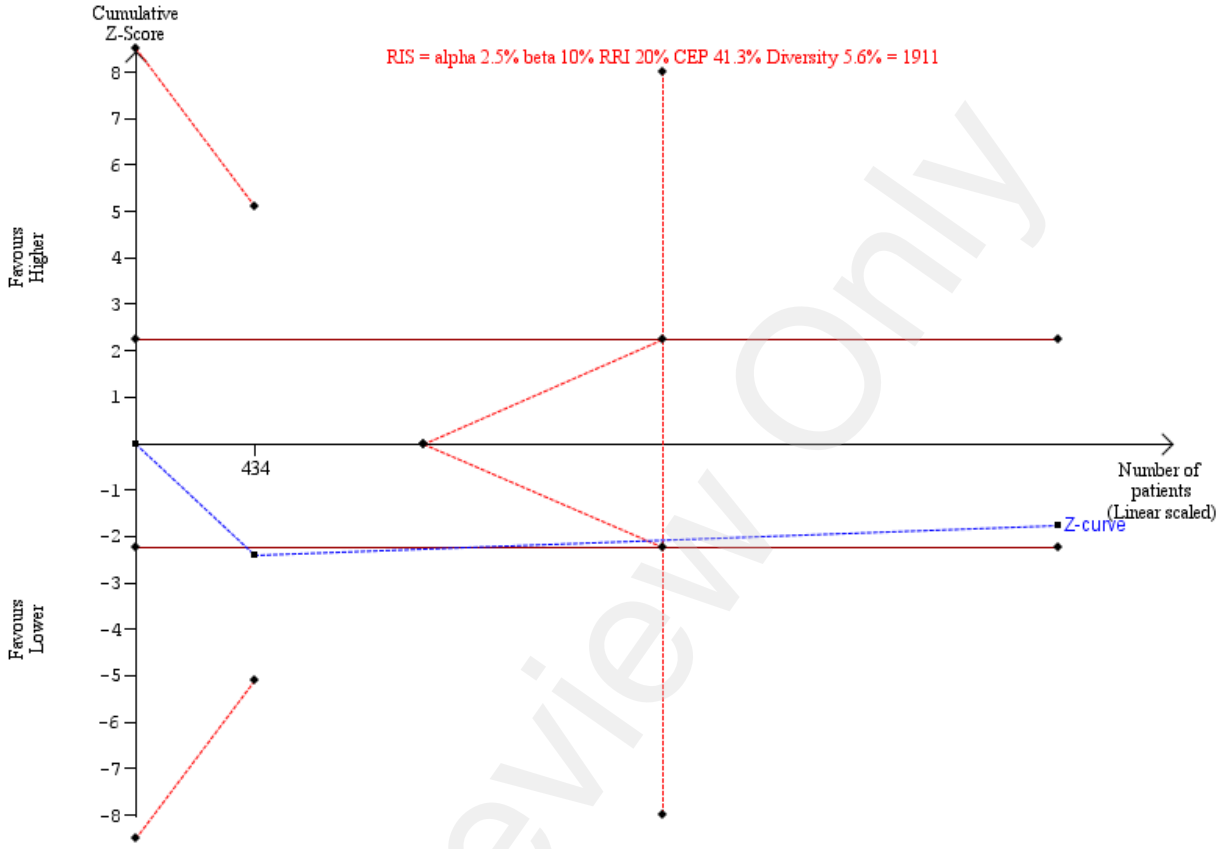
Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 27%; Chi^2 = 2.73; P = 0.14) indicated statistical heterogeneity for trials reporting on one or more SAE as a composite outcome.

For the highest reported proportion of any SAE (I^2 = 38%; Chi^2 = 20.94; P = 0.7) and for the cumulated number of SAEs (I^2 = 67%; Chi^2 = 30.37; P = 0.0002) we found significant heterogeneity.

Trial Sequential Analyses

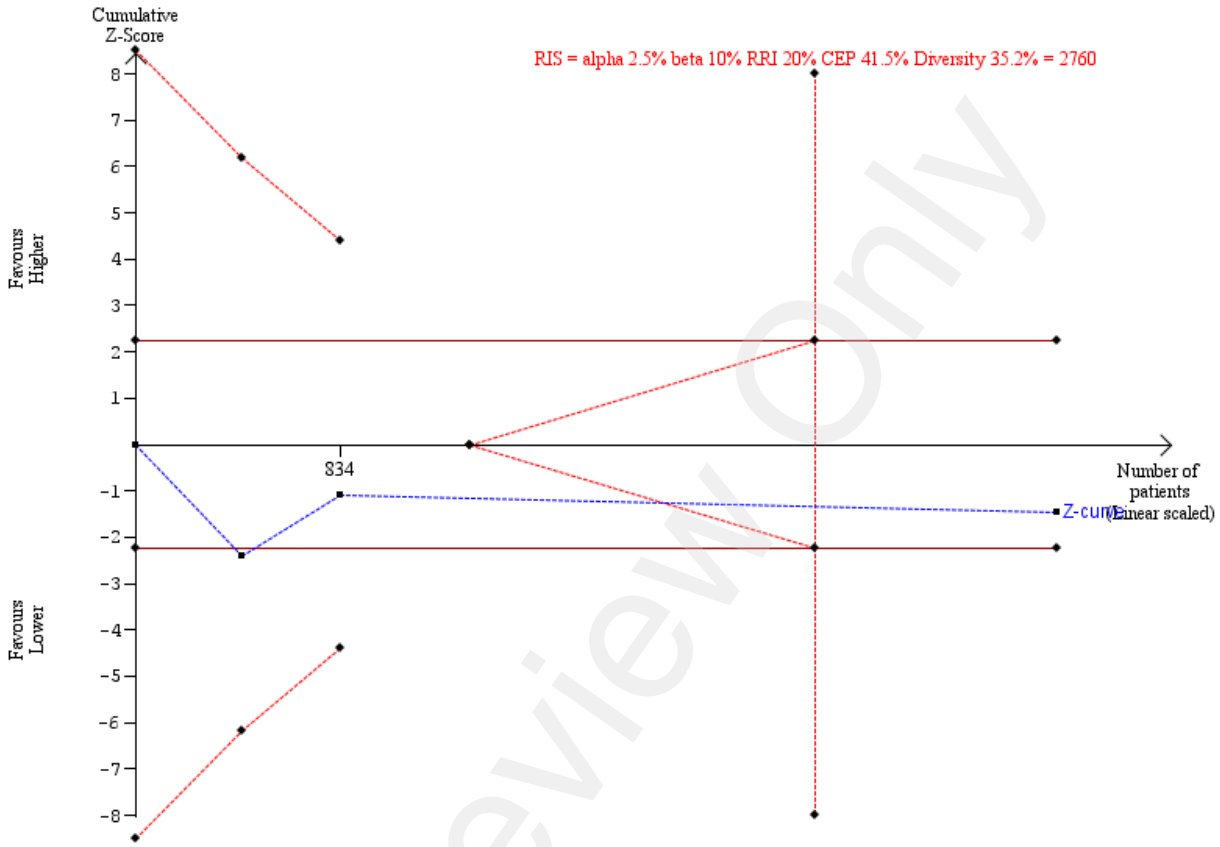
Trial Sequential Analysis of trials judged to be at overall low risk of bias of the proportion of participants with one or more SAE showed that with an anticipated RRI of 20%, proportion of participants with one or more SAE in the control group of 41.3%, a type 1 error level of 2.5%, a type two error level of 10%, and a diversity of 5.6%, the required information size was 1,911 participants. The cumulated Z-curve crossed the trial sequential monitoring boundaries for futility, and with 3,344 participants in the analysis the required information size was exceeded. This indicated that, considering repetitive testing, the evidence was sufficient to refute a 20% RRI or a 20% RRR for benefit or harm of higher versus lower oxygenation strategies (Figure 7). The TSA CI was 0.88 to 1.30.

Figure 7. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the proportion of participants with one or more serious adverse events in trials judged to be at overall low risk of bias. The analysis was based on a porportion of participants with one or more serious adverse events in the control group (control event proportion = CEP) of 41.3%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 5.6%. The cumulative Z-curve crossed the boundaries for trial sequential monitoring boundaries for futility, and the required information size (RIS) was exceeded.



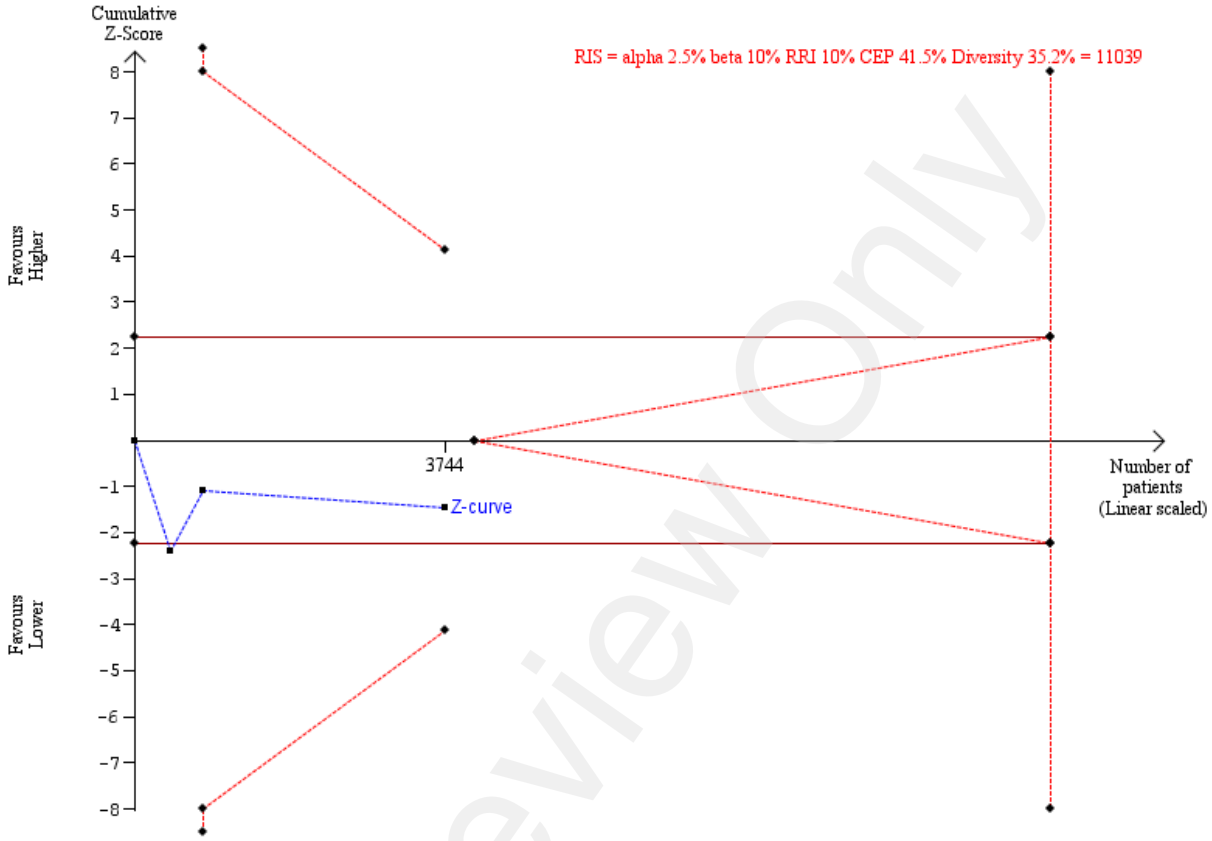
Trial Sequential Analysis of the proportion of participants with one or more SAE showed that with an anticipated RRI of 20%, proportion of participants with one or more SAE in the control group of 41.5%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 35.2%, the required information size was 2,760 participants. The cumulated Z-curve crossed the trial sequential monitoring boundaries for futility, and with 3,744 participants in the analysis the required information size was exceeded. This indicated that, considering repetitive testing, the evidence was sufficient to refute a 20% RRI or a 20% RRR for benefit or harm of higher versus lower oxygenation strategies (Figure 8). The TSA CI was 0.90 to 1.24.

Figure 8. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the proportion of participants with one or more serious adverse events. The analysis was based on a proportion of participants with one or more serious adverse events in the control group (control event proportion = CEP) of 41.5%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 35.2%. The cumulative Z-curve crossed the boundaries for trial sequential monitoring boundaries for futility, and the required information size (RIS) was exceeded.



Trial Sequential Analysis of the proportion of participants with one or more SAE with an anticipated RRI of 10%, proportion of participants with one or more SAE in the control group of 41.5%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 35.2%, the required information size was 10,944 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 10% RRR or a 10% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 9). The TSA CI was 0.91 to 1.23.

Figure 9. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the proportion of paticipants with one or more serious adverse events. The analysis was based on a proportion of participants with one or more serious adverse events in the control group (control event proportion = CEP) of 41.5%, a relative risk increase (RRI) of 10%, a type 1 error level (alpha) of 2.5%, a type 2 error level (beta) of 10%, and a diversity of 35.2%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of the analysis of the proportion of participants with one or more SAE closest to the null-effect (RRR 2%), proportion of patients with one or more SAE the control group of 41.5%, a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 35.2%, only 1.39% of the required information size was reached for this outcome. The required information size was 269,350 participants.

Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

No significant subgroup differences were found according to: overall risk of bias (Analysis 3.2); types of oxygen interventions (Analysis 3.3); level of FiO_2/target in higher group (Analysis 3.4); level of FiO_2/target in lower group (Analysis 3.5); or oxygen delivery system (Analysis 3.7). When considering effects according to ICU population only subtotals are presented as more than one trial is represented in more than one subgroup (Analysis 3.6).

Sensitivity analyses

The sensitivity analysis excluding trials comparing two low oxygenation strategies or two high oxygenation strategies indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies on the proportion of participants with one or more SAE (RR 1.07, 95% CI 0.99 to 1.15; I^2 = 6%; 3344 participants; 2 trials; Analysis 3.8).

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

- Best-worst-case scenario fixed-effect meta-analysis: RR 0.95, 95% CI 0.88 to 1.02; I^2 = 97%; 3944 participants; 3 trials; Analysis 3.9);
- Worst-best-case scenario fixed-effect meta-analysis: RR 1.19, 95% CI 1.11 to 1.27; I^2 = 93%; 3944 participants; 3 trials; Analysis 3.10).

Data were imputed for three trials (Asfar 2017; Gelissen 2021; Schjørring 2021).

Fourteen of 16 trials included (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Gomersall 2002; Ishii 2018; Jakkula 2018; Jun 2019; Lång 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019), with at total of 6349 participants randomized, reported on the occurrence of any SAE reported on outcomes categorized by us as an SAE according to the ICH-GCP definition (ICH-GCP 1997). Eight trial were judged to be at overall low risk of bias (Asfar 2017; Barrot 2020; Jakkula 2018; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019).

A list of SAEs only reported by one trial is provided in Table 1.

For sensitivity purposes, we estimated the reported proportion of participants with one or SAE in two ways:

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

Meta-analysis of the highest proportion of specific SAEs in trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (fixed-effect model RR 1.00, 95% CI 0.94 to 1.06; I^2 = 42%; 4945 participants; 8 trials; Analysis 1.3).

Meta-analysis of the cumulated number of SAEs in trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (fixed-effect model RR 1.02, 95% CI 0.98 to 1.06; I^2 = 0%; 4212 participants; 5 trials; Analysis 1.4).

Meta-analysis of all trials indicated no difference in effect of higher versus lower oxygenation strategies when assessing the highest reported proportion of specific SAEs in each trial (fixed-effect model RR 1.00, 95% CI 0.95 to 1.06; I^2 = 38%; 6031 participants; 14 trials; Analysis 3.11). Individual types of SAEs included mortality (Barrot 2020; Girardis 2016; Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Mazdeh 2015; Panwar 2016; Schjørring 2021; Yang 2019); proportion of participants with one or more SAE (Asfar 2017; Gelissen 2021); mechanical ventilation (reported as a poor outcome) (Gomersall 2002); and myocardial infarction (Jun 2019).

Meta-analysis of all trials indicated no evidence of a difference of higher versus lower oxygenation strategies when assessing the cumulated number of SAEs (fixed-effect model RR 1.03, 95% CI 1.00 to 1.06; I^2 = 67%; 6053 participants; 14 trials; Analysis 3.12).

Individual types of SAEs 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); intestinal ischaemia; coma; digestive haemorrhage; acute myocardial infarction; seizures; stroke; cardiac arrhythmia, hemodynamically instability; and unexplained brain oedema on computed tomography (CT) scan.

Quality of life

One included trial, with a total of 1000 participants randomized, reported data on quality of life (Mackle 2020). The trial was at overall high risk of bias for this outcome as data was only available for 499 of 617 eligible patients.

The mean (± SD) reported health state scores (EQ-VAS) at 180 days after randomization were 67.6 points (± 22.4) in the higher group (253 participants) versus 70.1 points (± 22.0) in the lower group (246 participants); mean difference -2.5 points (95% CI -6.4 to 1.4; P = 0.22; 499 participants; Summary of findings 2).

Publication bias

No evaluation of publication bias was performed as less than 10 trials reported on quality-of-life.

Heterogeneity

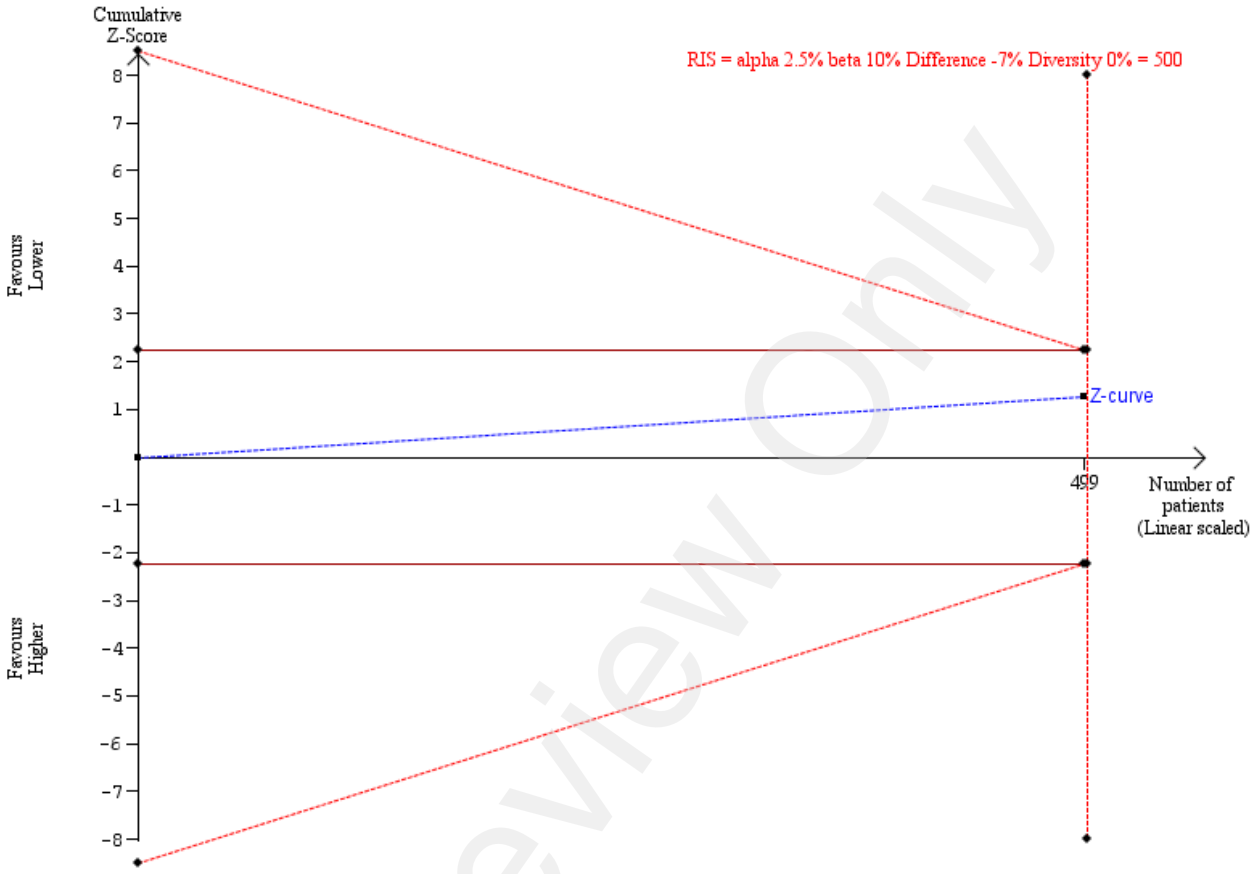
As only one included trial reported on quality of life, no estimation of heterogeneity could be performed.

Trial Sequential Analyses

The minimum important difference for the EQ-VAS score has been estimated to be between 7 and 12 (Patrona 2014; Pickard 2007; Zanini 2015). We chose the smallest effect size for reference of analysis.

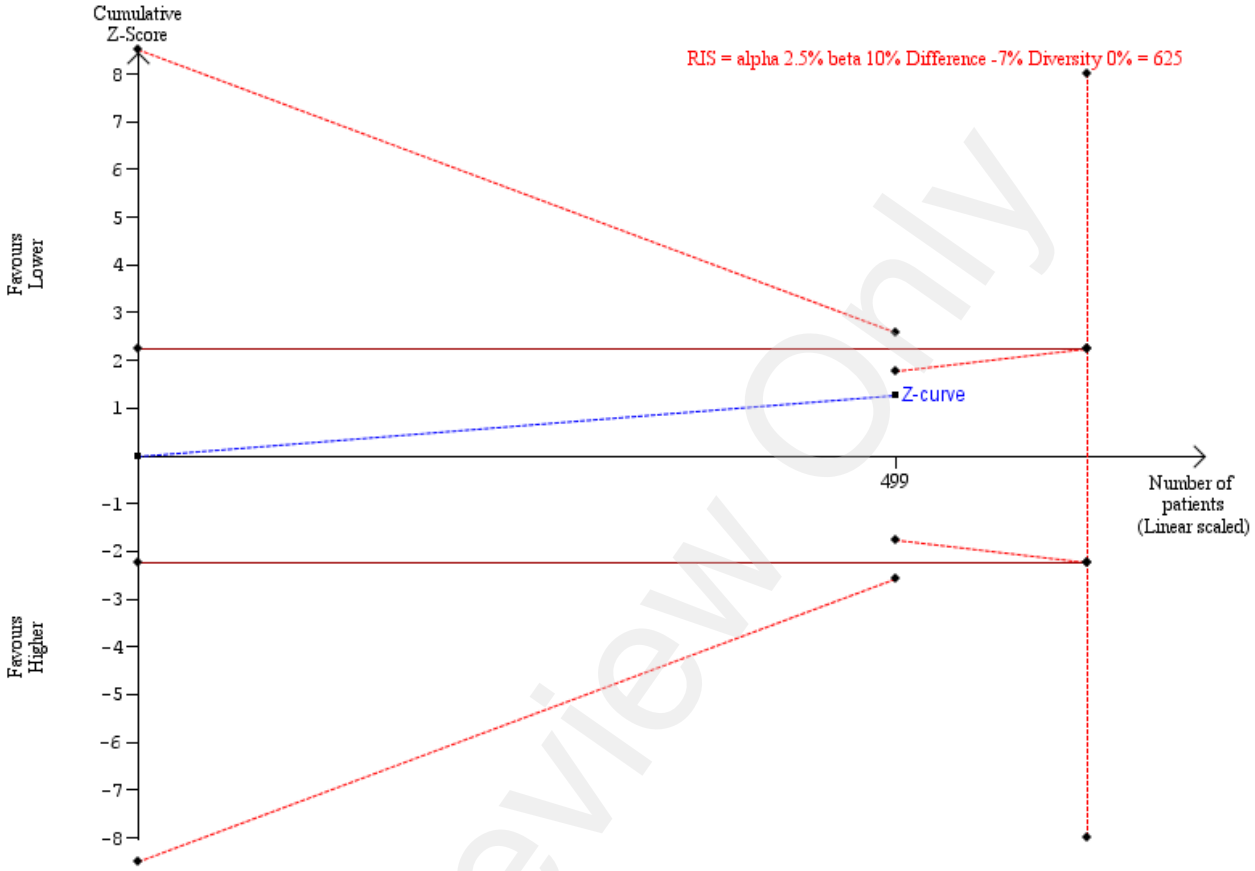
Trial Sequential Analysis of quality of life showed that with an anticipated mean difference of -7 points, a type 1 error level of 2.5%, a type 2 error level of 10%, and a diversity of 0%, the required information size was 500 participants. The cumulative Z-curve did not cross any boundaries for benefit 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 7 point increase or a 7 point decrease benefit or harm of higher versus lower oxygenation strategies (Figure 10). The TSA CI was -6.96 to 1.96. However, the analysis was only one participant short in order to obtain the required information size of 500.

Figure 10. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on quality of life. The analysis was based on a observed mean difference of -2.5 points, an minimal clinical relevance of -7 points, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



As the diversity for the Trial Sequential Analysis of the quality of life was 0%, we performed a sensitivity analysis with diversity defined as 20%, as described in the protocol (Barbateskovic 2017). With an anticipated mean difference of -7, a type 1 error level of 2.5%, a type 2 error level of 10%, and a diversity of 20%, the required information size was 625 participants. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 7 point increase or a 7 point decrease benefit or harm of higher versus lower oxygenation strategies (Figure 11). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was -7.59 to 2.59.

Figure 11. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on quality of life. The analysis was based on a observed mean difference of -2.5 points, an minimal clinical relevance of -7 points, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 20%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of the analysis of quality of life closest to the null-effect (1.4 point increase), a type 1 error of 2.5%, a type 2 error of 10%, and a diversity of 0%, only 4.0% of the required information size was reached for this outcome. The required information size was 12,484 participants.

Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

As only one included trial reported on quality of life (Mackle 2020), none of the preplanned subgroup analyses (Barbateskovic 2017) were performed.

Sensitivity analyses

As only one included trial reported on quality of life, no sensitivity analyses were performed.

Secondary outcomes

Lung injury

None of the 16 included trials reported any data on lung injury (as a composite outcome defined as either ARDS, pulmonary fibrosis, or pneumonia) at any time-point. Seven of the 16 trials reported on specific lung outcomes during index admission: 3 trials reported on ARDS (Gelissen 2021; Jakkula 2018; Panwar 2016); 3 trials reported on pneumonia (Asfar 2017; Girardis 2016; Mackle 2020), one trial reported on both ARDS and pneumonia (Lång 2018); no trials reported on pulmonary fibrosis.

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).

Three trials reporting on lung injury were judged to be at overall low risk of bias (432 participants) (Barrot 2020; Jakkula 2018; Panwar 2016).

Meta-analysis of trial at low risk of bias of the highest proportion of participants with lung injury as a composite outcome in trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (random-effects model RR 1.16, 95% CI 0.74 to 1.81; I^2 = 0%; 424 participants; 3 trials; Analysis 1.5; very low certainty evidence; Summary of findings 1).

Meta-analysis of trials at low risk of bias of the cumulated number of participants with lung injury as a composite outcome in trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (random-effects model RR 1.16, 95% CI 0.74 to 1.81; I^2 = 0%; 424 participants; 3 trials; Analysis 1.6; very low certainty evidence; Summary of findings 1).

Meta-analysis of all trials indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the highest reported proportion of lung injury (fixed-effect model RR 1.06, 95% CI 0.82 to 1.36; I^2 = 0%; 1942 participants; 7 trials; Analysis 4.1; very low certainty evidence; Summary of findings 2).

Meta-analysis af all trials indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the cumulated number of lung injury events (random-effects model RR 1.02, 95% CI 0.80 to 1.31; I^2 = 0%; 1942 participants; 7 trials; Analysis 4.2; very low certainty evidence; Summary of findings 2).

ARDS

Four of 16 trials, with a total of 871 participants randomized, reported on ARDS (Gelissen 2021; Jakkula 2018; Lång 2018; Panwar 2016). A total of 2.7% in the higher group versus 3.8% in the lower group had ARDS. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of ARDS (random-effects model RR 0.86, 95% CI 0.43 to 1.69; I^2 = 0%; 862 participants; 4 trials; Analysis 4.3; very low certainty evidence).

Pneumonia

Four of 16 trials, with a total of 1197 participants randomized, reported on pneumonia (Asfar 2017; Barrot 2020; Girardis 2016; Lång 2018). A total of 16.3% in the higher group versus 15.1% in the lower group had pneumonia. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of pneumonia (random-effect model RR 1.08, 95% CI 0.82 to 1.41; I^2 = 0%; 1145 participants; 4 trials; Analysis 4.4; very low certainty evidence).

Publication bias

No evaluation of publication bias was performed as less than 10 trials reported on lung injury (as a composite outcome defined as either ARDS, pulmonary fibrosis, or pneumonia) at any time-point.

Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor indicated statistical heterogeneity when assessing the highest reported number of lung injuries (I^2 = 0%; Chi^2 = 2.09; P = 0.91) or the cumulated number of lung injuries (I^2 = 0%; Chi^2 = 4.44; P = 0.62).

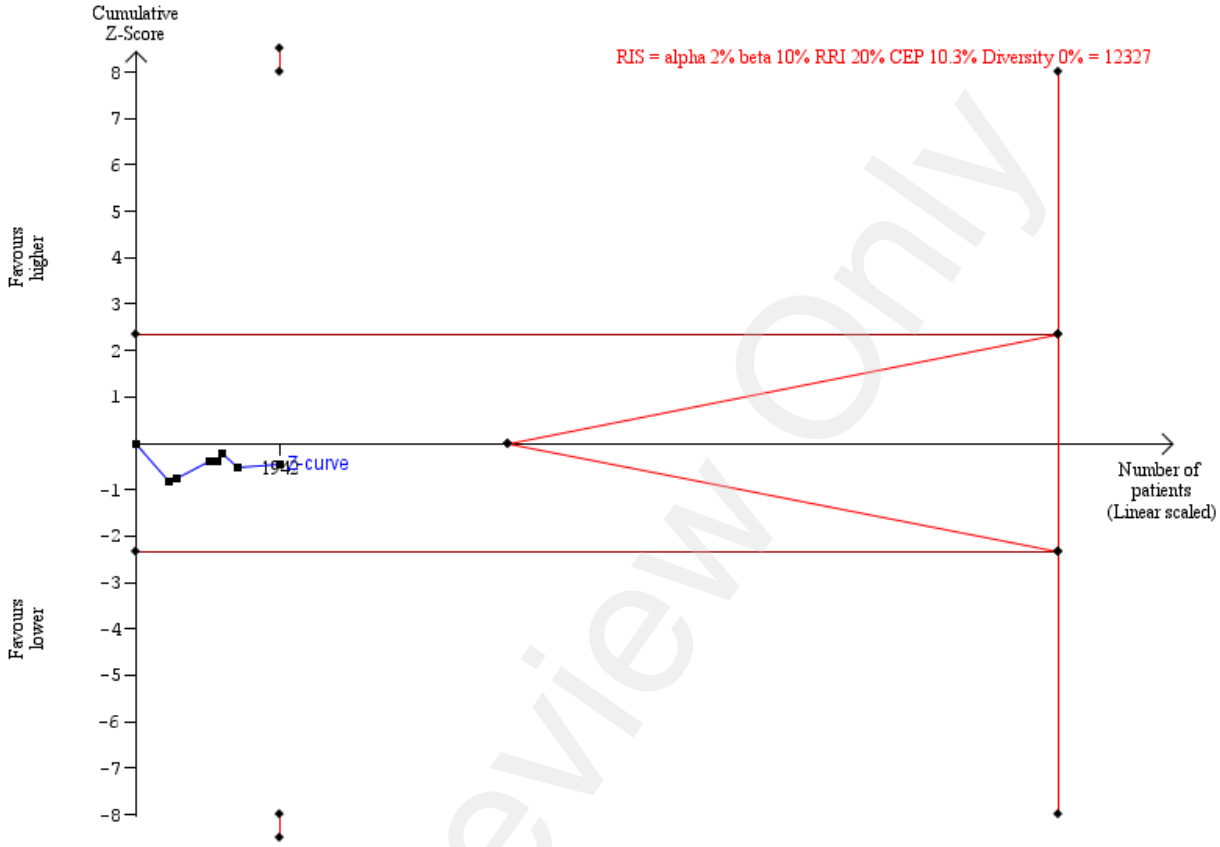
Trial Sequential Analyses

As no trial reported on lung injury as a composite outcome, no TSA was performed.

Highest reported proportion of lung injuries

Trial Sequential Analysis of all trials of the highest reported proportion of lung injuries showed that with an anticipated RRI of 20%, lung injury in the control group of 10.3%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, the required information size was 12,327 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% RRR or a 20% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 12). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.38 to 2.96.

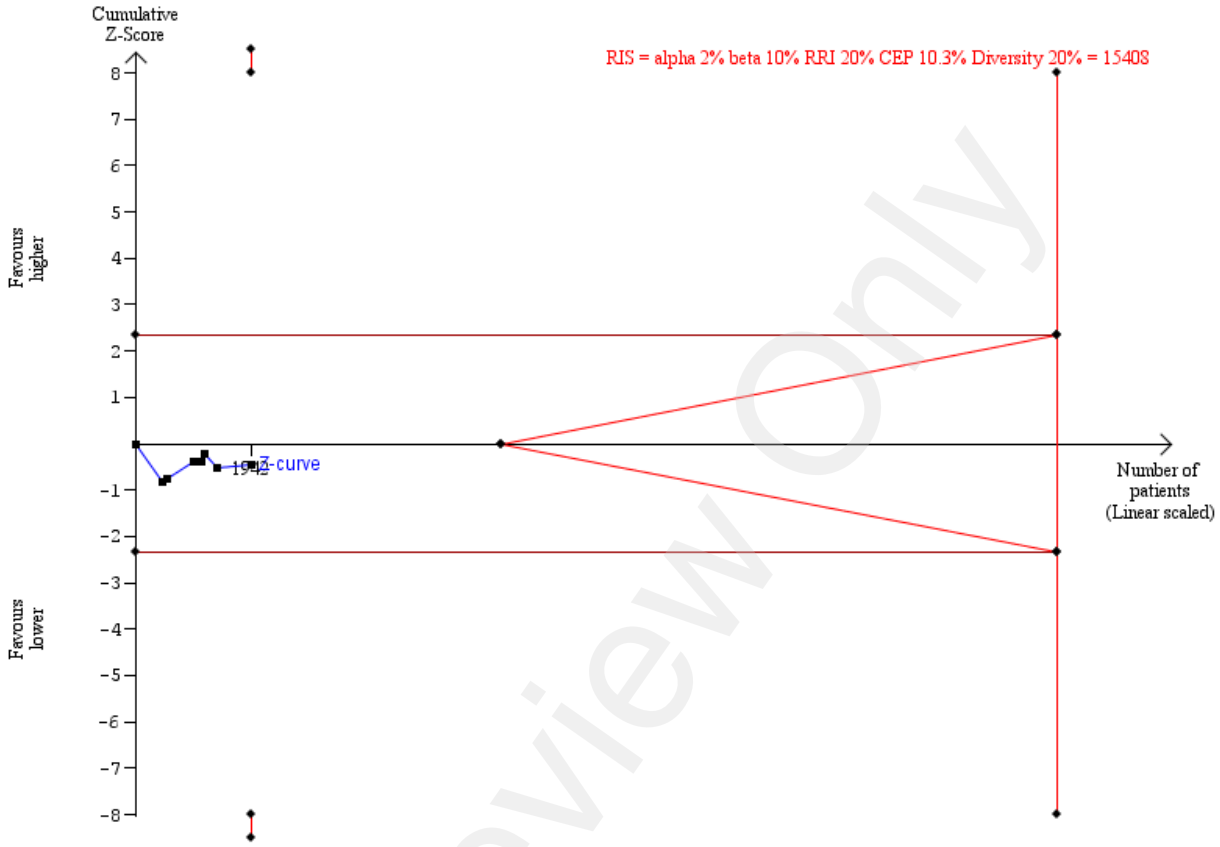
Figure 12. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the highest reported proportion of lung injuries. The analysis was based on a highest reported proportion of lung injuries in the control group (control event proportion = CEP) of 10.3%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Trial Sequential Analysis of all trials of the highest reported proportion of lung injuries with an anticipated RRI of 10%, lung injury in the control group of 10.3%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 4.10% of the required information size reached for this outcome. The required information size was 47,339 participants.

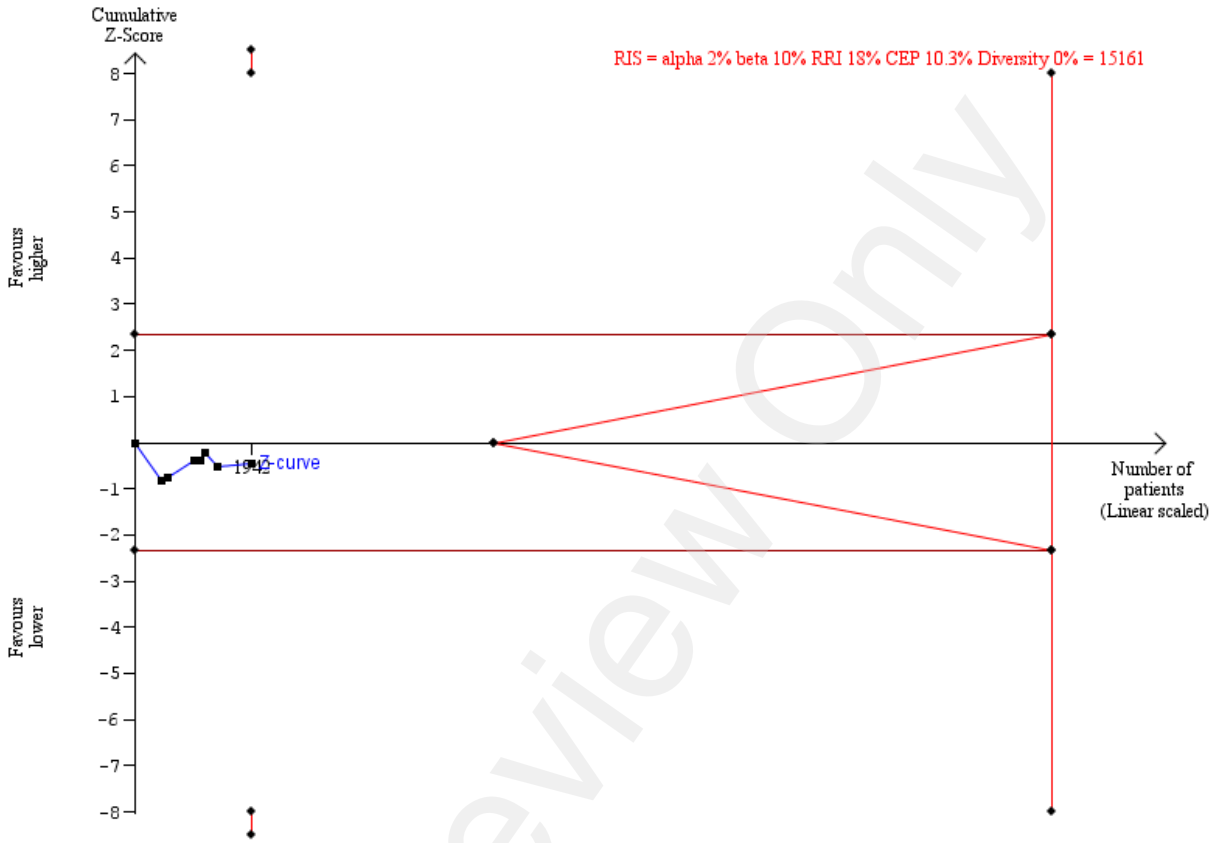
As the diversity for the Trial Sequential Analysis of the highest reported proportion of lung injuries was 0%, we performed a sensitivity analysis with diversity defined as 20%, described in the protocol (Barbateskovic 2017). With an anticipated RRI of 20%, lung injury in the control group of 10.3%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 20%, the required information size was 15,408 participants. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR or a 20% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 13). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.38 to 2.96.

Figure 13. Trial Sequential Analysis of the effects of higher versus oxygenation strategies on the highest reported proportion of lung injuries. The analysis was based on a highest reported proportion of lung injuries in the control group (control event proportion = CEP) of 10.3%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 20%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



For the Trial Sequential Analysis of all trials using an RRR or RRI based on the conventional 95% confidence interval of the analysis of the highest reported proportion of lung injuries closest to the null-effect (RRR of 18%), lung injury in the control group of 10.3%, a type 1 error level of 2%, a type two error level of 10%, and a diversity of 0%, the required information size was 15,161 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 an 18% RRR or an 18% RRI for benefit or harm of higher versus lower oxygenation strategies(Figure 14). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.38 to 2.94.

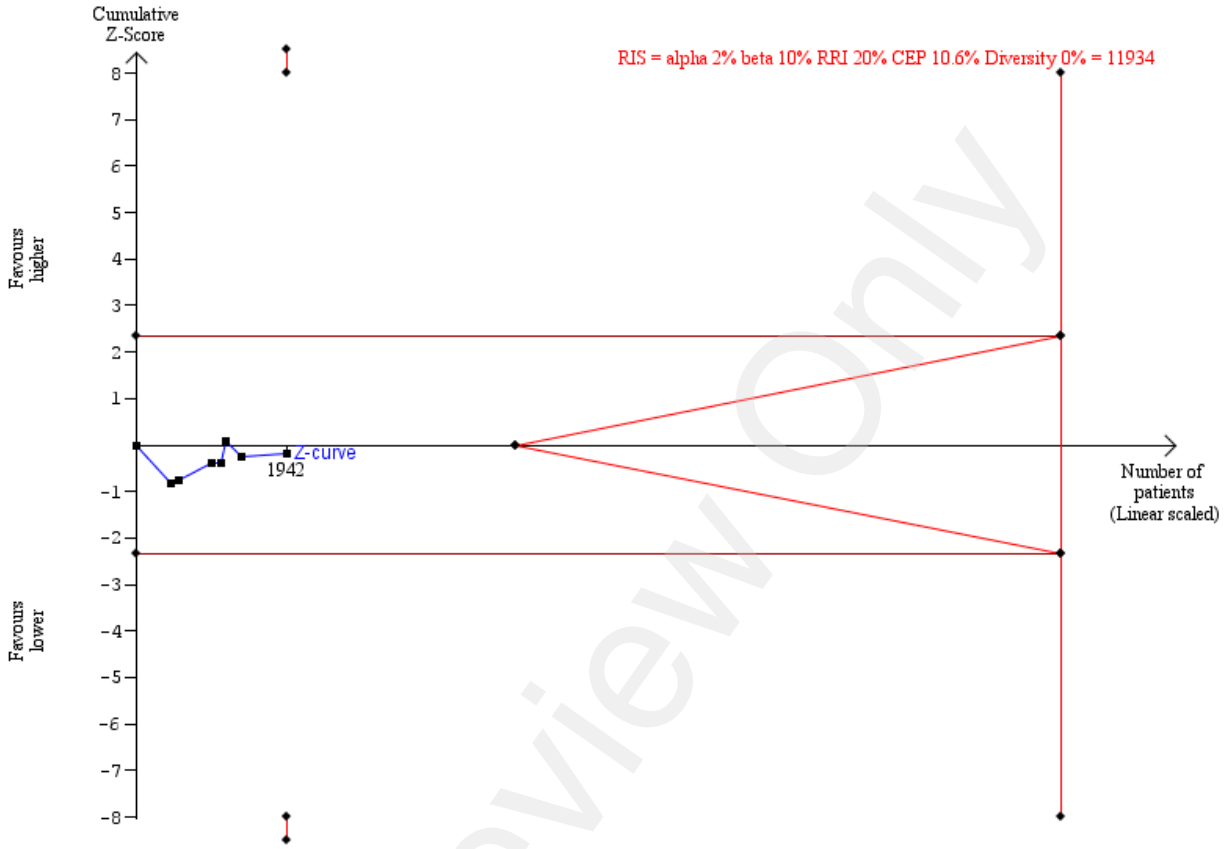
Figure 14. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the highest reported proportion of lung injuries. The analysis was based on a highest reported proportion of lung injuries in the control group (control event proportion = CEP) of 10.3%, a relative risk reduction (RRR) of 18%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Cumulated number of lung injuries

Trial Sequential Analysis of all trials of the cumulated number of lung injuries showed that with an anticipated RRI of 20%, lung injury in the control group of 10.6%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, the required information size was 11,934 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% RRR or a 20% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 15). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.31 to 3.44.

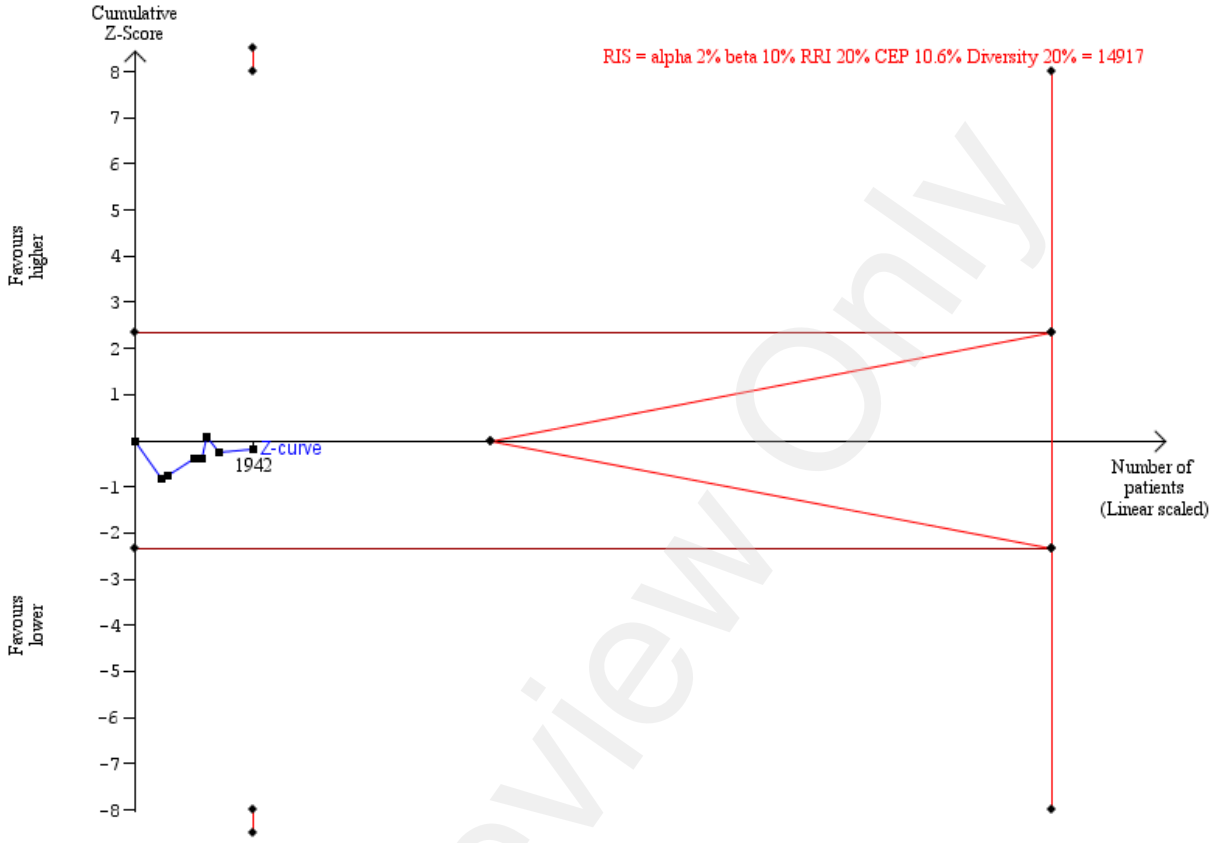
Figure 15. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the cumulated number of lung injuries. The analysis was based on a cumulated number (control event proportion = CEP) of 10.6%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Trial Sequential Analysis of all trials of the cumulated number of lung injuries with an anticipated RRI of 10%, lung injury in the control group of 10.3%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 4.24% of the required information size reached for this outcome. The required information size was 45,837 participants.

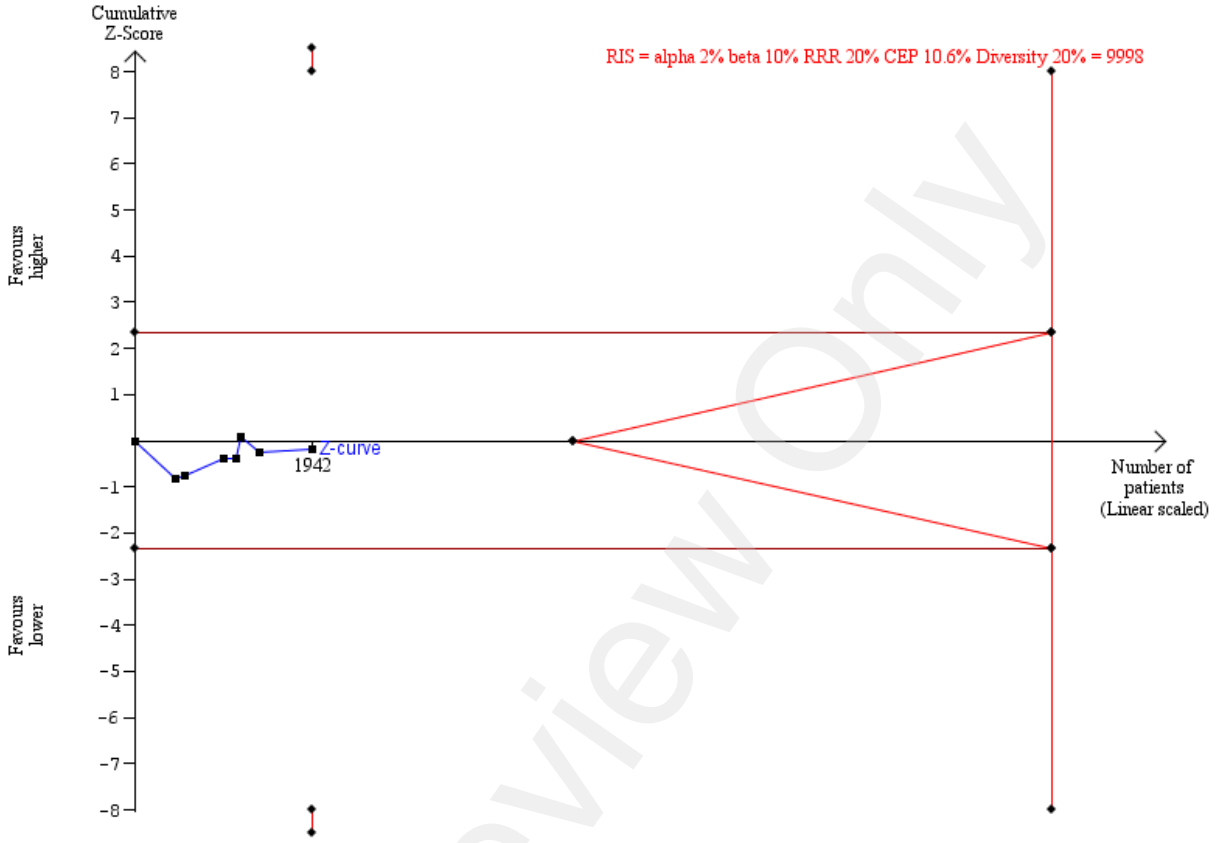
As the diversity for the Trial Sequential Analysis of the cumulated number of lung injuries was 0%, we performed a sensitivity analysis with diversity defined as 20%, as described in the protocol (Barbateskovic 2017). With an anticipated RRI of 20%, lung injury in the control group of 10.6%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 20%, the required information size was 14,917 participants. This indicated that considering sparse data and repetitive testing, evidence was insufficient to confirm or refute a 20% RRR or a 20% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 16). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.31 to 3.44.

Figure 16. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the cumulated number of lung injuries. The analysis was based on a cumulated number of lung injuries in the control group (control event proportion = CEP) of 10.6%, a relative risk increase (RRI) of 20%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 20%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of the cumulated number of lung injuries at longest follow-up closest to the null-effect (RRR of 20%), lung injury in the control group of 10.6%, a type 1 error level of 2%, a type two error level of 10%, and a diversity of 0%, the required information size was 9,998 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% RRR or a 20% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 17). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.31 to 3.44.

Figure 17. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the cumulated number of lung injuries. The analysis was based on a cumulated number of lung injuries in the control group (control event proportion = CEP) of 10.6%, a relative risk reduction (RRR) of 20%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

As none of the 16 included trials reported on lung injury (as a composite outcome defined as either ARDS, pulmonary fibrosis, or pneumonia) at any time-point, none of the prespecified subgroup analyses (Barbateskovic 2017) were performed.

Myocardial infarction

Three of 16 trials included, with a total of 3589 participants randomized, reported on the occurrence of myocardial infarction (Gelissen 2021; Jun 2019; Schjørring 2021). No sensitivity analysis was performed for acute myocardial infarction according to risk of bias, as only one trial reporting on this outcome was judged to be at overall low risk of bias (Schjørring 2021). In this trial 8/1457 participants in the higher group and 14/1453 in the lower group had an acute myocardial infarction (random-effects model RR 0.57, 95% CI 0.24 to 1.35; 2910 participants; very low certainty evidence; Summary of findings 1).

In all trials, a total of 1.0% in the higher group versus 1.7% in the lower group had myocardial infarction. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of myocardial infarction (random-effects model RR 0.59, 95% CI 0.25 to 1.38; I^2 = 17%; 3368 participants; 3 trials; Analysis 5.1; very low certainty evidence; Summary of findings 2).

Publication bias

No evaluation of publication bias was performed as less than 10 trials reported on acute myocardial infarction.

Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 44%; Chi^2 = 3.55; P = 0.17) indicated statistical heterogeneity for trials reporting on acute myocardial infarction.

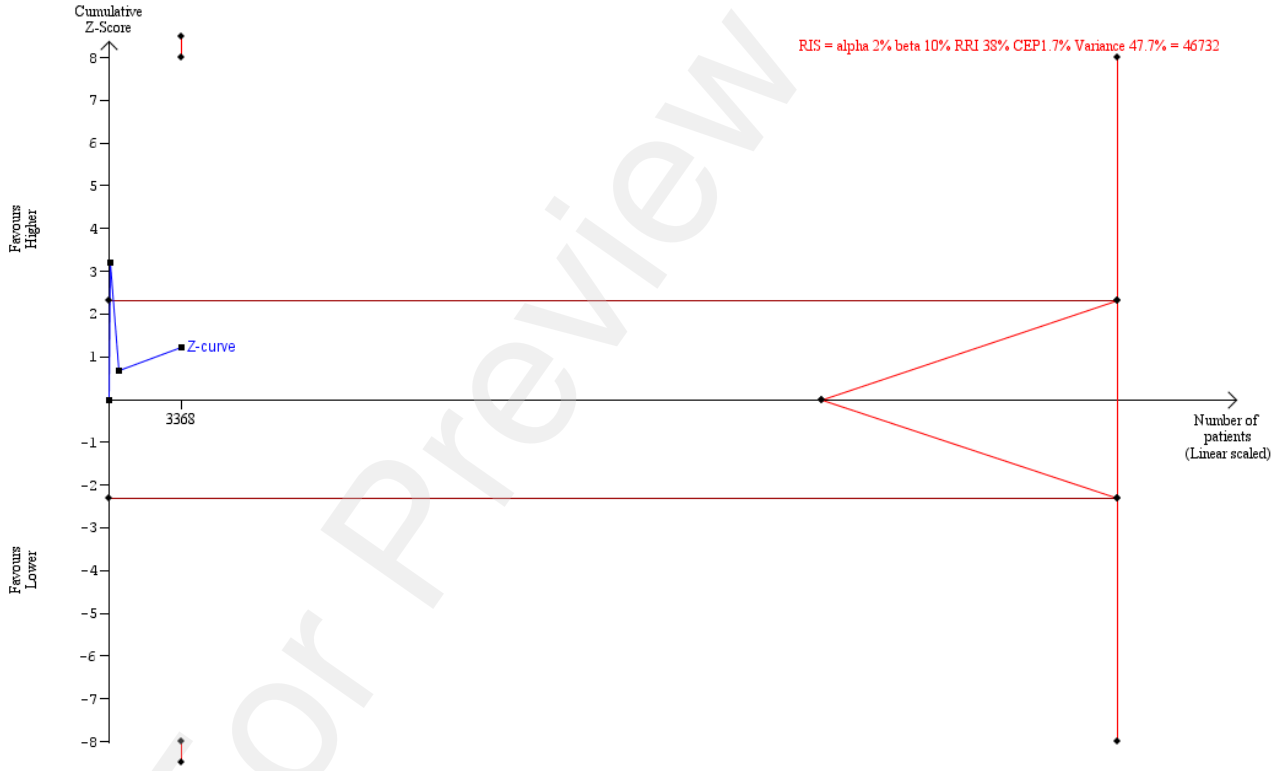
Trial Sequential Analyses

Trial Sequential Analysis of all trials reporting on the occurrence of myocardial infarction with an anticipated RRI of 20%, myocardial infarction in the control group of 1.7%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 47.7%, only 2.13% of the required information size reached for this outcome. The required information size was 157,971 participants.

Trial Sequential Analysis of all trials reporting on the occurrence of myocardial infarction with an anticipated RRI of 10%, myocardial infarction in the control group of 1.7%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 47.7%, only 0.56% of the required information size reached for this outcome. The required information size was 603,682 participants.

For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of all trials reporting on the occurrence of myocardial infarction closest to the null-effect (RRI of 38%), myocardial infarction in the control group of 1.7%, a type 1 error level of 2%, a type two error level of 10%, and a diversity of 47.7%, the required information size was 46,732 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 38% RRR or a 38% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 18). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.02 to 18.57.

Figure 18. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the occurrence of myocardial infarction. The analysis was based on myocardial infarction in the control group (control event proportion = CEP) of 1.7%, a relative risk reduction (RRI) of 38%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 47.7%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

No significant subgroup differences were found according to: overall risk of bias (Analysis 5.2); types of oxygen interventions (Analysis 5.3); level of FiO_2/target in lower group (Analysis 5.5); or oxygen delivery system (Analysis 5.7). No test for subgroup differences could be performed according to level of FiO_2/target in higher group (Analysis 5.4). When considering effects according to ICU population only subtotals are presented as more than one trial is represented in more than one subgroup (Analysis 5.6).

Sensitivity analyses:

No sensitivity analysis excluding trials comparing two low oxygenation strategies or two high oxygenation strategies was performed, as only one trial reporting on acute myocardial infarction satisfied the criteria for this analysis (Schjørring 2021).

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

- Best-worst-case scenario random-effects meta-analysis: RR 0.17, 95% CI 0.06 to 0.52; I^2 = 75%; 3551 participants; 3 trials; Analysis 5.8);
- Worst-best-case scenario random-effects meta-analysis: RR 1.72, 95% CI 0.17 to 17.40; I^2 = 95%; 3551 participants; 3 trials; Analysis 5.9).

Data were imputed two trials (Gelissen 2021; Schjørring 2021).

Stroke

Four of 16 trials included (4707 participants) reported on the occurrence of stroke (Barrot 2020; Gelissen 2021; Mackle 2020; Schjørring 2021). Two trials were judged to be at overall low risk of bias (3133 participants) (Barrot 2020; Schjørring 2021). Meta-analysis of trials at overall low risk of bias indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies (fixed-effect model RR 1.04, 95% CI 0.59 to 1.83; I^2 = 49%; 3111 participants; 2 trials; Analysis 1.7; very low certainty evidence; Summary of findings 1)

In all included trials, a total of 1.2% in the higher group versus 1.2% in the lower group had stroke. Meta-analysis of all trials showed no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of stroke (fixed-effect model RR 1.12, 95% CI 0.65 to 1.92; I^2 = 0; 4476 participants; 4 trials; Analysis 6.1; very low certainty evidence; Summary of findings 2).

Publication bias

No evaluation of publication bias was performed as less than 10 trials reported on the occurrence of stroke.

Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 0%; Chi^2 = 2.60; P = 0.46) indicated statistical heterogeneity for trials reporting on the occurrence of stroke.

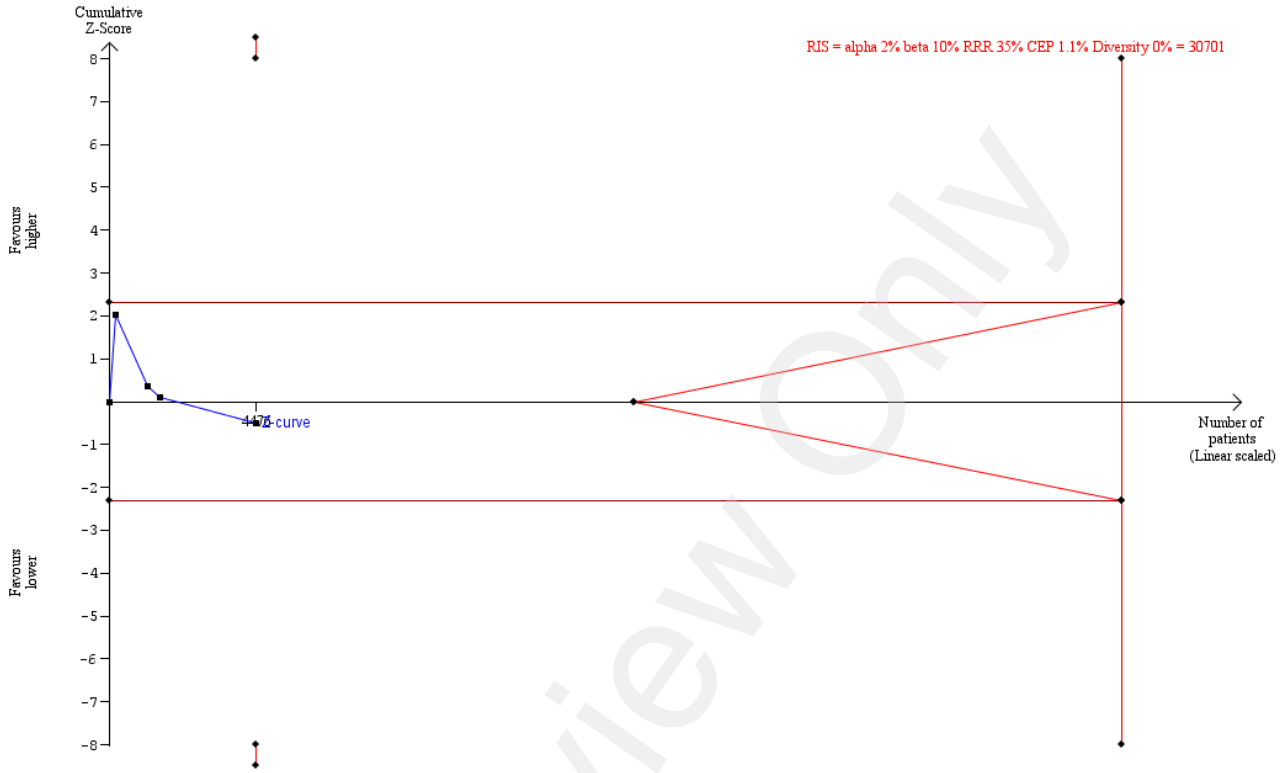
Trial Sequential Analyses

Trial Sequential Analysis of all trials reporting on the occurrence of stroke with an anticipated RRI of 20%, stroke in the control group of 1.1%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 3.48% of the required information size reached for this outcome. The required information size was 128,595 participants.

Trial Sequential Analysis of all trials reporting on the occurrence of stroke with an anticipated RRI of 10%, stroke in the control group of 1.1%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 0.91% of the required information size reached for this outcome. The required information size was 491,270 participants.

For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of all trials reporting on the occurrence of stroke closest to the null-effect (RRR of 35%), stroke in the control group of 1.1%, a type 1 error level of 2%, a type two error level of 10%, and a diversity of 0%, the required information size was 30,701 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 35% RRR or a 35% RRI for benefit or harm of higher versus lower oxygenation strategies (Figure 19). The TSA CI, adjusted for multiple outcomes, sparse data, and repetitive testing, for the intervention effect was 0.12 to 11.15.

Figure 19. Trial Sequential Analysis of the effects of higher versus lower oxygenation strategies on the occurrence of stroke. The analysis was based on stroke in the control group (control event proportion = CEP) of 1.1%, a relative risk reduction (RRR) of 35%, a type 1 error level (alpha) of 2.0%, a type 2 error level (beta) of 10%, and a diversity of 0%. Required information size = RIS. The cumulative Z-curve did not cross any boundaries for benefit and harm, nor trial sequential monitoring boundaries for futility.



Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

No significant subgroup differences were found according to: overall risk of bias (Analysis 6.2); level of FiO_2/target in lower group (Analysis 6.5); or oxygen delivery system (Analysis 6.7). No test of subgroup differences could be performed according to types of oxygen interventions (Analysis 6.3) or level of FiO_2/target in higher group (Analysis 6.4). When considering effects according to ICU population only subtotals are presented as one trial is represented in more than one subgroup (Analysis 6.6).

Sensitivity analyses

The sensitivity analysis excluding trials comparing two low oxygenation strategies or two high oxygenation strategies indicated no evidence of a difference in the effect of higher versus lower oxygenation strategies on the occurrence of stroke (RR 1.04, 95% CI 0.59 to 1.83; I^2 = 49%; 3111 participants; 2 trials; Analysis 6.8).

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

- Best-worst-case scenario fixed-effect meta-analysis: RR 0.19, 95% CI 0.13 to 0.29; I^2 = 92%; 4707 participants; 4 trials; Analysis 6.9);
- Worst-best-case scenario random-effects meta-analysis: RR 5.86, 95% CI 3.84 to 8.96; I^2 = 92%; 4707 participants; 4 trials; Analysis 6.10).

Data were imputed for four trials (Barrot 2020; Gelissen 2021; Mackle 2020; Schjørring 2021).

Sepsis

Two of 16 trials included, with at total of 685 participants randomized, reported on the occurrence of sepsis at any time-point (Barrot 2020; Girardis 2016). No sensitivity analysis was performed for sepsis according to risk of bias, as only one trial reporting on this outcome was judged to be at overall low risk of bias (Barrot 2020). In this trial 19/102 participants in the higher group experienced a new episode of sepsis, and 11/99 in the control group (fixed-effect RR 1.68, 95% 0.84 to 3.34, 201 participants; Summary of findings 1).

In all included trials, a total of 12.5% in the higher group versus 6.9% in the lower group had sepsis. Meta-analysis indicated evidence of benefit from lower oxygenation strategies compared with higher when assessing the occurrence of sepsis (random-effects model RR 1.81, 95% CI 1.11 to 2.95; I^2 = 0%; 646 participants; 2 trials; Analysis 7.1 ; very low certainty evidence; Summary of findings 2).

Publication bias

No evaluation of publication bias was performed as less than 10 trials reported on the occurrence of sepsis at any time-point.

Heterogeneity

Neither visual inspection of the forest plot nor inconsistency factor (I^2 = 0%; Chi^2 = 0.09; P = 0.76) indicated statistical heterogeneity for trials reporting on the occurrence of sepsis.

Trials Sequential Analyses

Trial Sequential Analysis of all trials reporting on the occurrence of sepsis with an anticipated RRI of 20%, sepsis in the control group of 6.9%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 3.37% of the required information size reached for this outcome. The required information size was 19,177 participants.

Trial Sequential Analysis of all trials reporting on the occurrence of sepsis with an anticipated RRI of 10%, sepsis in the control group of 6.9%, a type 1 error of 2%, a type 2 error of 10%, and a diversity of 0%, only 0.88% of the required information size reached for this outcome. The required information size was 73,494 participants.

For the Trial Sequential Analysis using an RRR or RRI based on the conventional 95% confidence interval of the analysis of the occurrence of sepsis closest to the null-effect (RRI of 11%), sepsis in the control group of 6.9%, a type 1 error level of 2%, a type two error level of 10%, and a diversity of 0%, only 1.18% of the required information sizes was reached. The required information size was 54,943 participants.

Bayes factor

Bayes factors for all outcomes are presented in Table 3.

Subgroup analyses

No significant subgroup differences were found according to: overall risk of bias (Analysis 7.2), or oxygen delivery system (Analysis 7.7). No test of subgroup differences could be performed according types of oxygen interventions (Analysis 7.3), level of FiO_2/target in higher group (Analysis 7.4), level of FiO_2/target in lower group (Analysis 7.5). When considering effects according to ICU population only subtotals are presented as no trial is represented in more than one subgroup, or according to ICU-population (Analysis 7.6).

Sensitivity analyses

No sensitivity analysis excluding trials comparing two low targets or two high targets was performed, as only one trial reporting on the occurrence of sepsis satisfied the criteria for this analysis (Barrot 2020).

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

- Best-worst-case scenario random-effects meta-analysis: RR 0.98, 95% CI 0.61 to 0.1.57; I^2 = 26%; 685 participants; 2 trials; Analysis 7.8);
- Worst-best-case scenario random-effects meta-analysis: RR 2.54, 95% CI 1.24 to 5.19; I^2 = 57%; 685 participants; 2 trials; Analysis 7.9).

Data were imputed for four trials (Barrot 2020; Girardis 2016).

Additional serious adverse events

Serious adverse events only reported in a single trial

SAEs only reported in a single trial are presented in Table 1.

Serious adverse events reported in two or more trials

The following SAEs were reported in two or more trials.

Delirium

Two of 16 trials included, with at total of 239 participants randomized, reported on the occurrence of delirium (Barrot 2020; Martin 2021). A total of 10.9% in the higher group versus 10.3% in the lower group had delirium. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of delirium (random-effects model RR 1.05, 95% CI 0.50 to 2.21; I^2 = 0%; 235 participants; 2 trials; Analysis 8.6).

Pneumothorax

Two of 16 trials included, with at total of 647 participants randomized, reported on pneumothorax (Asfar 2017; Barrot 2020). A total of 4.7% in the higher group versus 3.5% in the lower group had pneumothorax. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of pneumothorax (random-effects model RR 1.34, 95% CI 0.63 to 2.28; I^2 = 0%; 635 participants; 2 trials; Analysis 8.7).

Intestinal ischaemia

Three of 16 trials included, with at total of 3575 participants randomized, reported on intestinal ischaemia (Asfar 2017; Barrot

2020; Schjørring 2021). A total of 2.0% in the higher group versus 2.3% in the lower group had intestinal ischaemia. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of intestinal ischaemia (random-effects model RR 0.90, 95% CI 0.35 to 2.30; I^2 = 46%; 3545 participants; 3 trials; Analysis 8.8).

Cardiovascular failure including shock

Two of 16 trials included, with at total of 3408 participants randomized, reported on cardiovascular failure including shock (Girardis 2016; Schjørring 2021). A total of 3.2% in the higher group versus 3.0% in the lower group had cardiovascular failure including shock. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of cardiovascular failure including shock (fixed-effect model RR 1.09, 95% CI 0.98 to 1.20; I^2 = 85%; 3355 participants; 2 trials; Analysis 8.9).

Cardiac arrhythmia

Two of 16 trials included, with at total of 239 participants randomized, reported on cardiac arrhythmia (Barrot 2020; Martin 2021). A total of 1.8% in the higher group versus 2.4% in the lower group had cardiac arrhythmia. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of cardiac arrhythmia (random-effects model RR 0.78, 95% CI 0.48 to 1.28; I^2 = 0%; 235 participants; 2 trials; Analysis 8.10).

Liver failure

Two of 16 trials included, with at total of 1054 participants randomized, reported on liver failure (Gelissen 2021; Girardis 2016). A total of 4.0% in the higher group versus 2.5% in the lower group had liver failure. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of liver failure (random-effects model RR 1.22, 95% CI 0.10 to 12.98; I^2 = 85%; 845 participants; 2 trials; Analysis 8.11).

Renal failure

Three of 16 trials included, with at total of 1088 participants randomized, reported on renal failure (Gelissen 2021; Girardis 2016; Martin 2021). A total of 11.2% in the higher group versus 11.8% in the lower group had renal failure. Meta-analysis indicated no evidence of a difference between higher versus lower oxygenation strategies when assessing the occurrence of renal failure (random-effects model RR 0.96, 95% CI 0.67 to 1.38; I^2 = 0%; 879 participants; 3 trials; Analysis 8.12).

DISCUSSION

Summary of main results

We included 16 trials that randomized a total of 6486 participants in this updated systematic review, of which 14 trials with a total of 6349 randomized participants contributed with data to the analyses. 13 trials reported on mortality; SAEs were reported in 14 trials, however only three of these trials reported the proportion of participants with one or more SAE as a composite outcome; 1 trial reported on quality of life; No trial reported on lung injury as a composite outcome (ARDS, pneumonia, or lung fibrosis), but 7 trials reported on the individual components; 3 trials reported on acute myocardial infarction; 4 trials reported on stroke; and 2 trials reported on severe sepsis.

Overall, we found no evidence for a beneficial or harmful effect of higher versus lower oxygenation strategies for adults admitted to the ICU, however the level of certainty for all outcomes were low or very low.

Analysis of all-cause mortality at maximum follow-up showed no difference between higher versus lower oxygenation strategies for neither trials judged to being at overall low risk of bias, nor when considering all included trials. Trial Sequential Analysis, considering multiple outcomes, sparse data, and repetitive testing, revealed that an RRI of 20% could be refused as the trial sequential monitoring border for futility was crossed (Figure 4).

Analysis of the proportion of participants with one or more SAE at maximum follow-up showed no difference between higher versus lower oxygenation strategies when considering trials at overall low risk of bias, nor when considering all included trials. Similar was found in the sensitivity analyses considering the highest reported proportion or cumulated number of SAEs. Trial Sequential Analysis, considering multiple outcomes, sparse data, and repetitive testing, revealed that an RRI of 20% of the proportion of participants with one or more SAEs at maximum follow-up could be refused as the trial sequential monitoring border for futility was crossed (Figure 8).

Only one trial reported on quality of life at any time post-randomization, with no indications of a difference between higher versus lower oxygenation strategies.

There was no evidence of a difference in the occurrence of lung injury at maximum follow-up with higher versus lower oxygenation strategies when analysed as the highest proportion, the cumulated number, or as individual components. However, the evidence is very uncertain (Analysis 4.1; Analysis 4.2; Analysis 4.3; Analysis 4.4) due to serious risk of bias and imprecision. Additionally, Trial Sequential Analysis considering multiple outcomes, sparse data, and repetitive testing, revealed that only 16% 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, or futility had been crossed (Figure 12).

We found no evidence of a difference in the occurrence of myocardial infarction at maximum follow-up when comparing higher versus lower oxygenation strategies in all included trials. Only one of three trials reporting on acute myocardial infarction was judged to be at overall low risk of bias.

We found no evidence of a difference in the occurrence of stroke at maximum follow-up when comparing higher versus lower oxygenation strategies in trials judged to be at overall low risk of bias, nor in all included trials.

Two trials reported on the occurrence of sepsis at maximum follow-up, indicating potential harm from higher versus lower oxygenation strategies, but the certainty of evidence was very low. Only one of the two trials reporting on severe sepsis was judged to be at overall low risk of bias for this outcome.

Overall completeness and applicability of evidence

In this updated review we reran the literature search and included RTCs up until April 2021 and identified seven additional trials (Barrot 2020; Gelissen 2021; Jun 2019; Mackle 2020; Martin 2021; Schjørring 2021; Yang 2019) in addition to the 10 trials ((Asfar 2017; Girardis 2016; Gomersall 2002; Ishii 2018; Jakkula 2018; Lång 2018; Mazdeh 2015; Panwar 2016; Taher 2016; Young 2017) included up to December 2018 in the previous version of this review (Barbateskovic 2019). One trial previously identiffied was excluded due to overlap in trial population with a new report (Young 2017).

We identified a high risk of clinical diversity/heterogeneity among the identified trials. See Characteristics of included studies. This was especially evident in relation to the applied interventions, as the trials did not use the same definitions of higher and lower oxygenation strategies. Some trials used a fixed FiO_2, whilst others used oxygenation targets of SpO_2, PaO_2, or SaO_2, or combinations of such. Thus, the achieved oxygen level may end up being high when compared to other trials, even though participants were allocated to the lower FiO_2 or oxygenation target group; only 6 of 16 trials assessed strategies categorized by us as true higher versus true lower (Asfar 2017; Barrot 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019). This is reflected in the sensitivity analyses excluding high versus high strategies and low versus low strategies (Analysis 2.8; Analysis 3.8; Analysis 6.8). Additionally, some trials were designed with an overlap between interventions (Girardis 2016; Gomersall 2002; Jun 2019; Mackle 2020) thus challenging inferences on outcome effects. Also, the duration of the interventions varied substantially, ranging from a few hours (Taher 2016) to the entirety of the ICU admission for up to 90 days including any readmissions (Schjørring 2021). A summary of interventions used in the included trials is presented in Table 2.

Clinical diversity/heterogeneity in the studied populations and trial settings was present (Barbateskovic 2021 (a)). The identified trials were conducted from 1994 (Gomersall 2002) to 2020 (Schjørring 2021). Eight trials were conducted in Europe (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Jakkula 2018; Lång 2018; Martin 2021; Schjørring 2021); two in China (Jun 2019; Yang 2019); two in Iran (Mazdeh 2015; Taher 2016); one in Australia and New Zealand (Mackle 2020); one in Australia, New Zealand, and France (Panwar 2016); one in Hong Kong (Gomersall 2002); and one in Japan (Ishii 2018).

Mean or median age of participants spanned from 44 years (Lång 2018) to 70 years (Schjørring 2021), and the percentage of males spanned from 49% (Jakkula 2018) to 84% (Lång 2018). The severity of disease was reported in various manners; one trial reported Simplified Acute Physiology Score (SAPS) II score (Girardis 2016); two trials reported SAPS III score (Asfar 2017; Barrot 2020); four trials reported Sequential Organ Failure Assessment (SOFA) score of which three reported the total score (Barrot 2020; Martin 2021; Schjørring 2021) and one trial reported SOFA scores excluding the respiratory component (Gelissen 2021); six trials reported Acute Physiology And Chronic Health Evaluation (APACHE) II score (Jakkula 2018; Lång 2018; Mackle 2020; Martin 2021; Panwar 2016; Yang 2019); two trials reported Glasgow Coma Scale (GCS) (Lång 2018; Taher 2016); one trial reported Barthel Index (Mazdeh 2015); and three trials did not report any score (Gomersall 2002; Ishii 2018; Jun 2019). Moreover, the severity of disease differed also among the included RCTs as illustrated by the median APACHE II score spanning from 17 (Yang 2019) to 28 (Jakkula 2018).

All RCTs' participants were admitted to an ICU prior to randomisation, but the inclusion criteria differed substantially; ten RCTs included multidisciplinary ICU patients (Asfar 2017; Barrot 2020; Gelissen 2021; Girardis 2016; Jun 2019; Mackle 2020; Martin 2021; Panwar 2016; Schjørring 2021; Yang 2019); eleven trials included patients regardless of gas exchange impairments of which six trials only included participant receiving mechanical ventilation (Asfar 2017; Ishii 2018; Lång 2018; Mackle 2020; Panwar 2016; Taher 2016) and four had no specific requirements for mechanical ventilation (Gelissen 2021; Girardis 2016; Jakkula 2018; Mazdeh 2015; Yang 2019); five RCTs included only participants with respiratory failure, but with varying definitions of such (Barrot 2020; Gomersall 2002; Martin 2021; Jun 2019; Schjørring 2021); three RCTs included only medical ICU patients (Gomersall 2002; Jakkula 2018; Mazdeh 2015); three RCTs included only surgical ICU patients (Ishii 2018; Lång 2018; Taher 2016); two RCTs included only participants with traumatic brain injury (Lång 2018; Taher 2016); one RCT included only participants with COPD (Gomersall 2002), one additional RCT allowed patients with COPD to be randomised (Schjørring 2021), whilst six RTCs excluded patients with COPD (Barrot 2020; Gelissen 2021; Girardis 2016; Jakkula 2018; Mackle 2020; Yang 2019); one RCT included only adults with acute stroke (Mazdeh 2015); and one RCT included only adults resuscitated from out-of-hospital cardiac arrest (Jakkula 2018). These obvious differences in inclusion criteria may have caused potentially different effect sizes of the applied interventions. Consequently, the overall generalisability of the findings to all types of ICU patients may be impeded. Statistical heterogeneity however, was evaluated as being low or moderate, except when considering the estimated cumulated number of participants with one or more serious adverse events, and number of participants with a new episode of cardiovascular failure including shock.

Missing or incomplete outcome data could potentially influence the effect estimates for several outcomes. Our sensitivity analysis on missing data (best-worst-case scenario and worst-best-case scenario; Analysis 2.9; Analysis 2.10) suggested that incomplete outcome data alone potentially could influence the results on all-cause mortality. The best-worst scenario (Analysis 2.9) indicated a benefit for higher oxygenation strategies, whilst the worst-best scenario (Analysis 2.10) indicated harm from higher oxygention strategies; however, both analyses were had CIs which included one (i.e. no difference in effect). Similar findings were made when considering the proportion of participants with one or more serious adverse events (Analysis 3.9; Analysis 3.10). When considering myocardial infarction, the sensitivity analyses suggested a significant impact of missing outcome data, at least in the best-worst-case scenario (Analysis 5.8). No significant effect was suggested in the worst-best-case scenario (Analysis 5.9). Missing outcome data for stroke was highly suggestive for a statistically significant impact on the effect estimate in the sensitivity analyses, as both the best-worst-case and worst-best-case scenarios produced statistically significant results in opposing directions (Analysis 6.9; Analysis 6.10). When considering sepsis,

the worst-best-case scenario suggested a significantly increased harm from higher oxygenation strategies (Analysis 7.9), but no significant effect was suggested in the best-worst-case scenario (Analysis 7.8).

The Trial Sequential Analyses on all-cause mortality and the proportion of participants with one or more SAEs could both reject a 20% RRR or RRI. Trial Sequential Analysis on SAEs revealed that the boundaries for futility was crossed in both analyses (Figure 4; Figure 8). We could also reject a 10% RRR or RRR for all-cause mortality. The TSA on lung injuries revealed that the required information size to detect or reject an RRR or RRI of 20% was not reached (Figure 12). This was also the case for myocardial infarction, stroke, and sepsis. However, further data for all outcomes are needed to establish more precise effect estimates.

Quality of the evidence

We used GRADEpro GDT to assess the certainty of the evidence for the results on all-cause mortality, SAEs, quality of life, lung injury, acute myocardial infarction, stroke, and severe sepsis both in trials judged to be at overall low risk of bias (Summary of findings 1) and in all included trials (Summary of findings 2).

We found no indications of publication bias when considering all-cause mortality, the estimated highest reported proportion of SAEs, or the estimated cumulated number of SAEs. For all other outcomes we were unable to assess publication bias due to limited data.

The GRADE assessment showed that the certainty of evidence for an effect on mortality in trials at overall low risk of bias was low due to indirectness. The certainty was very low when considering all trials included due to risk of bias and indirectness. Trial Sequential Analysis showed that the trial sequential monitoring borders for futility were crossed. Thus, even when considering repetitive testing and disregarding risk of bias, the evidence was sufficient to refute a 20% RRI or RRR for benefit or harm from higher versus lower oxygenation strategies.

The certainty of the evidence for effect on the proportion of SAEs in trials at overall low risk of bias was low due to indirectness and very low when considering all trials due to the added risk of bias. Trial Sequential Analysis showed that the boundary for futility was crossed and the required information size was reached; hence, evidence was sufficient to reject an RRI of at least 20%.

The certainty of evidence was very low for quality of life due to inconsistency, risk of bias, and imprecision. Only one trial was identified, and this trial was at overall low high of bias for this outcome.

The certainty of the evidence was very low for lung injury in both trials at low risk of bias due to indirectness, and imprecision, but also in all identified trials due to risk of bias, indirectness, and imprecision.

The certainty of the evidence was very low for acute myocardial infarction, stroke, and severe sepsis in both trials at low risk of bias due to indirectness, and imprecision, but also in all identified trials due to risk of bias, indirectness, and imprecision.

Potential biases in the review process

Strengths

For this updated review we used a predefined (Barbateskovic 2017), rigorous, up-to-date systematic review methodology, and made only minor changes to the methods during the updating process (Differences between protocol and review). In addition, we expanded the search strategy and screened a vast volume of reports.

As is now recommended, we used the RoB-2 tool to evaluate the risk of bias (systematic errors) for all outcomes (Higgins 2016; Sterne 2019), and GRADEpro GDT to assess the certainty of evidence. We complemented the conventional meta-analyses with Trial Sequential Analyses, incorporating adjusted thresholds for significance as a sensitivity analyses and as means to rigorously control the risk of random errors. To further investigate the magnitude of effects we post-hoc performed Trial Sequential Analyses assessing an RRI or RRR of 10%. Additionally, we used the eight-step procedure as suggested by Jakobsen 2014a when assessing if the thresholds or statistical and clinical significance were crossed.

We have provided detailed summaries of all included trials, as well as of all identified ongoing trials.

We post-hoc increased the power in the meta-analyses from 80% to 90% as meta-analyses should use the same or higher power as the included trials, in order to communicate the best available evidence.

Our findings were also supplemented by sensitivity analyses to assess the robustness.

Trials were included regardless of publication type, publication status, language, and outcomes reported. We contacted relevant trial authors if additional information was required.

We conducted two post-hoc analyses that estimated the effects of higher versus lower oxygenations strategies on the risk of having one or more SAE, and lung injury, respectively.

We performed analysis of publication bias on all-cause mortality, and detected no indications of publication bias for this estimate.

Limitations

To the best of our knowledge, we have identified all relevant trials for this review. Screening of reports has been carried out in two rounds with different authors. TLK and FMN did the screening for the update and were assisted by the primary authors OLS and MB.

Five authors (TLK, OLS, JW, AP and BSR) were investigators of one of the included trials (Schjørring 2021). Assessment of this trial and data extraction was validated by another author (FMN), not involved in this trial, to avoid bias in the process.

We identified 10 ongoing trials (ACTRN12620000391976; ChiCTR-INR-17012800; ChiCTR-IOR-17011717; CTRI/2020/12/029614; ISRCTN13384956; NCT02999932; NCT03141099; NCT04198077; NCT04425031; NCT04824703) comparing higher versus lower

oxygenation strategies in adult ICU patients. Publication of these trials will increase the precision of the estimates of effects.

We have made several post-hoc changes to the analyses pre-planned in the protocol to improve the quality of the final updated review. All changes are listed in the section Differences between protocol and review.

When evaluating risk of bias we judged several trials to be at overall low risk of bias, for several outcomes, despite none of the identified trials were fully blinded (i.e. blinding of participants/relatives, staff, and outcome assessors). Even though data should be included from trials where blinding of participant and/or personnel is not possible (Pocock 2015), inadequate blinding may still represent a limitation since it has been associated with underestimation of adverse intervention effects and amplifications of positive effects (Hrobjartsson 2014; Savovic 2018). On this basis, a biased effect estimate can not be precluded based on the included trials. However, as according to the updated RoB-2 tool (Sterne 2019) we did not downgrade the certainty of the evidence on mortality on this basis.

We analysed the individual components of 'serious adverse events', because each component of a composite outcome may not be equally severe and therefore could distort the result of the outcome (Garattini 2016). If more serious adverse events occur in one group than in the other, there is a risk of ignoring the actual differences in severity when a composite outcome is used. We identified three trials reporting on serious adverse events as a composite outcome of 'participants with one or more serious adverse events' (Asfar 2017; Gelissen 2021; Schjørring 2021). As we did not identify any trials that reported on the composite outcome of 'lung injury', we analysed each component separately here as well. To supplement the analysis of the effect on the composite outcomes of 'serious adverse events' and 'lung injury' reported in the included trials, we chose to update the post-hoc defined analyses on this matter; highest reported proportion and cumulated proportion. When analysing the 'highest reported proportion of serious adverse events' and 'highest reported proportion of lung injuries' the analyses included both participants from the event with the highest proportion, but the same participants may also have had other serious adverse events or lung injuries. If mortality, for example, was the highest proportion of SAEs, the analysis would imply that all participants that survived did not experience another SAEs. Thus, the analysis necessarily underestimates the 'proportion of participants with one or more SAE', as it would be reasonable to expect participants who are not included in the highest reported proportion to experience other SAEs not included in the highest reported proportion. On the contrary, the analyses ascertaining the 'cumulated proportion of SAEs' and 'cumulated number of lung injuries' inadvertently risk overestimating the intervention effect, as it implies that all participants who experience an SAE or lung injury only had this one specific event. Again, it is reasonable to expect at least one participant to experience more than one SAE. Thus, the 'true' effect is expected to reside in between these two extremes.

Agreements and disagreements with other studies or reviews

Several systematic reviews of RCTs on oxygenation strategies in critically ill patients have been published in recent years. For this review, we included only trials assessing adults admitted to and randomized within the ICU and without limiting findings to specific patients categories, whereas other reviews also included other settings, e.g. ARDS, trauma, surgery, or prehospitally initiated oxygen supplementation (Cabello 2016; Chu 2018; Cumpstey 2020; Eskesen 2018; Sepehrvand 2018; You 2018; Barbateskovic 2021 (a); Li 2021 (a); Hansen 2021). Meta-analyses report conflicting findings, with some reporting that higher levels of supplemental oxygen may be harmful or not beneficial (Cabello 2016; Chu 2018; Sepehrvand 2018; You 2018), whilst others report insufficient evidence to support beneficial or harmful effects of higher versus lower oxygenation strategies (Barbateskovic 2021 (a); Cumpstey 2020; Eskesen 2018; Li 2021 (a); Li 2021 (b); Hansen 2021). However, only few of these meta-analyses included proper risk of bias assessment to support their conclusions/recommendations. Additionally, we did not find the available evidence to be of high certainty, as has previously been suggested (Chu 2018). Higher versus lower oxygations strategies in patients with ARDS has also been evaluated in a recent Cochrane Review (Cumpstey 2020), but the authors only identified a single trial on this matter (Barrot 2020), thus making inferences difficult. Young et al. also investigated conservative versus liberal oxygen therapy in the ICU in patients with cardiact arrest in an individual-level data meta analysis, indicating benefit of a conservative oxygenation strategy (Young 2021). Limitations due to clinical diversity/heterogeneity are to a greater or lesser extent highlighted in these reviews, but seems only to be partly reflected in the conclusions. For this review we have in greater detail summarised the differences in the included trials and incorporated some of these aspects in our subgroup and sensitivity analyses. Most importantly the applied interventions in the identified trials varied substantially, thus making direct comparison of the trials difficult. This factor could also potentially contribute to imprecision in the overall effect estimate.

Despite low estimates of statistical heterogeneity for most of the effect estimates presented in this review, such estimates may not adequately account for the potentially extensive clinical diversity/heterogeneity. A novel tool for systematically assessing clinical diversity/heterogeneity in meta-analyses of intervention has been proposed (Barbateskovic 2021 (b)). By systematically identifying and quantifying domains of clinical diversity/heterogeneity the authors of systematic reviews and meta-analyses may incorporate these levels of information to conduct for example meta-regression analyses, that allows for adjustment according to important imbalances between identified trials. However, this was not performed as it was beyond the scope of this review.

We performed Trial Sequential Analyses 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 oxygenation strategies. This was also used by Chu and colleagues but without adjusting for multiple outcomes and using a potentially inadequate power of 80% (Chu 2018). Trial Sequential Analyses with adjustment for multiple outcomes and with proper power were conducted by Barbateskovic and colleagues when considering harms or benefits of higher or lower oxygenation strategies in a broader population of acutely ill adult patients (Barbateskovic 2021 (a)) equivalent of the population in the review by Chu et al. (Chu 2018).

Despite methodological discrepancies between our review and other meta-analyses and reviews, we agree with recently published reviews that the amount of data on this matter is still insufficient.

However, based on the available data, large intervention effects from oxygen therapy appear unlikely, as we could reject an overall RRI or RRR of 10% for mortality and 20% for SAEs in the ICU population. Correspondingly, Barbateskovic and colleagues could reject an RRI or RRR of 15% for mortality and 20% for SAEs in acutely ill patients (Barbateskovic 2021 (a)). However, even smaller effect sizes are relevant to patients given the widespread use of oxygen supplementation in the ICU.

Currently, we did not find that the presently available evidence necessitates a clinical practice guideline recommending a specific target or range of FiO_2, SpO_2, SaO_2, or PaO_2. This is particularly due to the very high clinical diversity/heterogeneity in the types of interventions, and durations of such, in the trials included in this review (Rasmussen 2018; Siemieniuk 2018).

AUTHORS' CONCLUSIONS

Implications for practice

The effects of higher versus lower oxygenation strategies for adults admitted to the intensive care unit on all-cause mortality, SAEs, quality of life, lung injuries, acute myocardial infarction, stroke, and severe sepsis at maximum follow-up, as defined by trialists, are still unclear due to low or very low certainty evidence. However, we could reject a relative risk increase or reduction of 10% on mortality and 20% on the proportion of patients with one or more SAE in the Trial Sequential Analyses of the included trials. Also, our results suggest a potential benefit of lower oxygenation strategies in relation to the occurrence of new sepsis. However, the certainty of the evidence was very low.

Implications for research

Randomized controlled trials assessing the benefits and harms of higher versus lower oxygen supplementation strategies are still warranted. Such trials should be conducted with the lowest possible risk of bias, lowest risk of other design errors, and lowest risk of random errors.

Oxygen is, in most countries world wide, considered a medical drug, thus it must be prescribed to patients balancing potential harmful and beneficial effects. The assessed interventions and durations of such should reflect clinically relevant and accepted stategeis of oxygen supplementation. Future trials should aim to differentiate the intervention groups so that trials are in fact comparing higher versus lower oxygenation levels. Furthermore, future trials should also focus on identifying those patients who in fact require oxygen supplementation.

Future trials should focus their assessments on multidisciplinary ICUs and critically ill adults in general, but should also incorporate stratification for important baseline risk factors that subsequently allows for testing of differences in outcomes in such groups. If possible, stratification according to presence or absence of hypoxaemia at baseline should also be considered.

As only few trials have reported outcome data beyond 90 days, extended follow-up periods, e.g. up to or beyond one year, should be considered to provide information on the long-term effects of higher versus lower oxygenation strategies. Reporting of core outcome sets, relevant to patients, should also be implemented.

ACKNOWLEDGEMENTS

We would like to thank (Content Editor), (Statistical Editor), (Peer Reviewers), (Consumer Referee), Janne Vendt (Information Specialist), and Vernon Paul Hedge (Managing Editors), and (Co-ordinating Editor) for their help and editorial advice during the preparation of this updated 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. [PMID: 28219612]

Barrot 2020 *{published data only}*

* Barrot L, Asfar P, Mauny F, Winiszewski H, Montini F, Badie J, et al. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. *New England Journal of Medicine* 2020;**382**(11):999-1008. [DOI: 10.1056/NEJMoa1916431] [PMID: 32160661]

Gelissen 2021 *{published data only}***10.1001/jama.2021.13011**

* Gelissen H, de Grooth HJ, Smulders Y, Wils E, de Ruijter W, Vink R, et al. Effect of low-normal vs high-normal oxygenation targets on organ dysfunction in critically ill patients: a randomized clinical trial. *JAMA* 2021;**326**(10):940-8. [DOI: 10.1001/jama.2021.13011] [PMID: 34463696]

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. [DOI: 10.1001/jama.2016.11993] [PMID: 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. [DOI: 10.1097/00003246-200201000-00018] [PMID: 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. In: Critical Care Medicine. Vol. 46. 2018:533. [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. [DOI: 10.1007/s00134-018-5453-9] [PMID: 30430209]

Jun 2019 *{published data only}***https://doi.org/10.1016/j.chest.2019.08.886**

* Jun J, Sun L, Wang Y, Liu F, Yang G, Han B. Invasive mechanical ventilation with high concentration oxygen therapy for AECOPD patients with acute myocardial infarction. In: Chest. Vol. 156. 2019 :Suppl A958. [DOI: 10.1016/j.chest.2019.08.886]

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. [DOI: 10.1111/aas.13093] [PMID: 29464691]

Mackle 2020 *{published data only}*

* The 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, et al. Conservative oxygen therapy during mechanical ventilation in the ICU. *New England Journal of Medicine* 2020 ;**382**(11):989-98. [DOI: https://doi.org/10.1056/NEJMoa1903297] [PMID: 31613432]

Martin 2021 *{published data only}***10.1177/17511437211010031**

* Martin DS, McNeil M, Brew-Graves C, Filipe H, O'Driscoll R, Stevens JL, et al. A feasibility randomised controlled trial of targeted oxygen therapy in mechanically ventilated critically ill patients. *Journal of the Intensive Care Society* 2021 ;**22**(4):280-7. [DOI: 10.1177/17511437211010031]

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. [PMID: 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. [DOI: 10.1164/rccm.201505-1019OC] [PMID: 26334785]

Schjørring 2021 *{published data only}*

Schjørring OL, Klitgaard TL, Perner A, Wetterslev J, Lange T, Siegemund M et al. Lower or higher oxygenation targets for acute hypoxemic respiratory failure. *New England Journal of Medicine* 2021 ;**384**(14):1301-11. [DOI: 10.1056/NEJMoa2032510] [PMID: 33471452]

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. [DOI: 10.5812/traumamon.26772] [PMID: 27218057]

Yang 2019 *{published data only}*

* Yang X, Shang Y, and YuanS. Low versus high pulse oxygen saturation directed oxygen therapy in critically ill patients: a randomized controlled pilot study. *Journal of Thoracic Disease* 2019 ;**11**(10):4234-40. [DOI: 10.21037/jtd.2019.09.66] [PMID: 10.21037/jtd.2019.09.66]

References to studies excluded from this review

Ahimahalle 2019 *{published data only}*

Ahimahalle TZ, Amirfarhangi A, Jabbari M, Jenabi A, Bagherzadegan H, Noghabaei G. Impact of oxygen therapy to ameliorate contrast-induced nephropathy in patients with acute coronary syndrome undergoing emergency angiography; a double-blinded clinical trial. *Journal of Renal Injury Prevention* 2019 ;**8**(4):283-8. [DOI: 10.15171/jrip.2019.52]

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. [PMID: 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. [PMID: 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. [PMID: 20959284]

Baekgaard 2019 *{published data only}*

Baekgaard JS, Isbye D, Ottosen CI, Larsen MH, Andersen JH, Rasmussen LS, et al. Restrictive vs liberal oxygen for trauma patients-the TRAUMOX1 pilot randomised clinical trial. *Acta Anaesthesiol Scandinavica* 2019;**63**(7):947-55. [DOI: 10.1111/aas.13362] [PMID: 30908592]

Bardsley 2018 *{published data only}*

Bardsley G, Pilcher J, McKinstry S, Shirtcliffe P, Berry J, Fingleton J, et al. Oxygen versus air-driven nebulisers for exacerbations of chronic obstructive pulmonary disease: a randomised controlled trial. *BMC Pulmonary Medicine* 2018 ;**18**(1):157. [DOI: 10.1186/s12890-018-0720-7] [PMID: 30285695]

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. [PMID: 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. [PMID: 29684433]

Butler 1987 *{published data only}*

Butler CM, Ham RO, Lafferty K, Cotton LT, Roberts VC. The effect of adjuvant oxygen therapy on transcutaneous pO2 and healing in the below-knee amputee. *Prosthetics and Orthotics International* 1987;**11**(1):10-16.

Cheng 2021 *{published data only}*

Cheng Z, Geng X, Tong Y, Dornbos D, Hussain M, Rajah GB, et al. Adjuvant high-flow normobaric oxygen after mechanical thrombectomy for anterior circulation stroke: a randomized clinical trial. *Neurotherapeutics* 2021 ;**18**(2):188-97. [DOI: 10.1007/s13311-020-00979-3] [PMID: 33410112]

Heidari 2017 *{published data only}*

Heidari F, Rahzani K, Iranpoor D, Rezaee K. The effect of oxygen on the outcomes of non-ST-segment elevation acute coronary syndromes. *IJC Metabolic and Endocrine* 2017 ;**14**:67-71. [DOI: 10.1016/j.ijcme.2016.12.002]

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. [PMID: 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 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. [DOI: 10.1111/echo.13599] [PMID: 28664557]

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. [PMID: 27893526]

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. [PMID: 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. [PMID: 19826023]

Mokhtari 2020 *{published data only}*

Mokhtari A, Akbarzadeh M, Sparv D, Bhiladvala P, Arheden H, Erlinge D, et al. Oxygen therapy in patients with ST elevation myocardial infarction based on the culprit vessel: results from

the randomized controlled SOCCER trial. *BMC Emergency Medicine* 2020;**20**(1):12. [DOI: 10.1186/s12873-020-00309-y]

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. [PMID: 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. [PMID: 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. [PMID: 22305833]

Rawles 1976 *{published data only}*

Rawles JM, Kenmure AC. Controlled trial of oxygen in uncomplicated myocardial infarction. *BMJ* 1976;**1**(6018):1121-3. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 28973619]

Sepehrvand 2019 *{published data only}*

Sepehrvand N, Alemayehu W, Rowe BH, McAlister FA, van Diepen S, Stickland M, et al. High vs. low oxygen therapy in patients with acute heart failure: HiLo-HF pilot trial. *ESC Heart Failure* 2019 ;**6**(4):667-77. [DOI: 10.1002/ehf2.12448] [PMID: 31102328]

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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 25261605]

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. [PMID: 29202261]

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. [PMID: 23957447]

References to ongoing studies

ACTRN12620000391976 *{published data only}*

ACTRN12620000391976. The Mega Randomised Registry Trial Comparing Conservative vs. Liberal OXygenation Targets (MEGA-ROX) https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=379432&isReview=true . who.int/trialsearch/Trial2.aspx?TrialID=ACTRN12620000391976 (first recieved: 10 March 2020).

ChiCTR-INR-17012800 *{published data only}*

ChiCTR-INR-17012800. The effect of conservative oxygen therapy in the mechanical ventilation patients. chictr.org.cn/showproj.aspx?proj=21892 (first recieved: 26 September 2017).

ChiCTR-IOR-17011717 *{published data only}*

ChiCTR-IOR-17011717. The effects of conservative oxygen therapy vs conventional oxygen therapy on outcomes in critically ill patients: a randomized controlled trial. chictr.org.cn/showproj.aspx?proj=19990 2017.

CTRI/2020/12/029614 *{published data only}*

CTRI/2020/12/029614. Liberal use of oxygen in early stage of COVID-19 patients. ctri.nic.in/Clinicaltrials/pdf_generate.php?trialid=49680&EncHid=&modid=&compid=%27,%2749680det%27 (first received 2 November 2020).

ISRCTN13384956 *{published data only}*

ISRCTN13384956. Intensive care unit randomised trial comparing two approaches to oxygen therapy (UK-ROX) . who.int/trialsearch/Trial2.aspx?TrialID=ISRCTN13384956 (first recieved: 8 December 2020).

NCT02999932 *{published data only}*

NCT02999932. Pulse Oxygen Saturation (SpO2) Directed Oxygen Therapy (POSDOT). clinicaltrials.gov/ct2/show/NCT02999932 (first recieved 14 December 2016).

NCT03141099 *{published data only}*

NCT03141099. Blood pressure and OXygenation targets after OHCA (BOX). clinicaltrials.gov/ct2/show/NCT03141099 (first received 4 May 2017).

NCT04198077 *{published data only}*

NCT04198077. Conservative Versus Conventional Oxygen Administration in Critically Ill Patients. clinicaltrials.gov/ct2/show/NCT04198077 (first recieved: 10 December 2019).

NCT04425031 *{published data only}*

NCT04425031 . Handling Oxygenation Targets in COVID-19 (HOT-COVID) . clinicaltrials.gov/ct2/show/NCT04425031 (first received June 5, 2020).

NCT04824703 *{published data only}*

NCT04824703. Comparative Study Between Liberal and Conservative Oxygen Therapy in Mechanically Ventilated Intensive Care Patients. . clinicaltrials.gov/ct2/show/NCT04824703 (first received 26 March, 2021).

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. [PMID: 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. anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12615000957594 (first received 15 August 2006).

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. [DOI: 10.1016/S0140-6736(10)60446-1] [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 12571129]

Barbateskovic 2021 (a)

Barbateskovic M, Schjørring OL, Krauss SR, Meyhoff CS, Jakobsen JC, Rasmussen BS, et al. Higher vs lower oxygenation strategies in acutely ill adults: a systematic review with meta-

analysis and trial sequential analysis. *Chest* 2021;**159**(1):154-73. [DOI: 10.1016/j.chest.2020.07.015] [PMID: 32687907]

Barbateskovic 2021 (b)

Barbateskovic M, Koster TM, Eck RJ, Maagaard M, Afshari A, Blokzijl F et al. A new tool to assess Clinical Diversity In Meta-analyses (CDIM) of interventions. *Journal of Clinical Epidemiology* 2021 ;**135**:29-41. [DOI: 10.1016/j.jclinepi.2021.01.023] [PMID: 33561529]

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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. Art. No: CD007160. [DOI: 10.1002/14651858.CD007160.pub4] [PMID: 27991651]

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. [PMID: 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. [PMID: 29726345]

Crapo 1986

Crapo JD. Morphologic changes in pulmonary oxygen toxicity. *Annual Review of Physiology* 1986;**48**:721-31. [PMID: 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. [PMID: 10556115]

Cumpstey 2020

Cumpstey AF, Oldman AH, Smith AF, Martin D, Grocott MP. Oxygen targets in the intensive care unit during mechanical ventilation for acute respiratory distress syndrome: a rapid review.. *Cochrane Database of Systematic Reviews* 2020 September 1, Issue 9. Art. No: CD013708. [DOI: 10.1002/14651858.CD013708] [PMID: 32870512]

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. [DOI: 10.1111/aas.12528.] [PMID: 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. [PMID: 25532567]

Deeks 2010

Deeks JJ, Higgins JPT. Statistical algorithms in Review Manager 5. https://training.cochrane.org/handbook/statistical-methods-revman5 2010:1-11.

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. [PMID: 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. [PMID: 19077208]

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. [PMID: 23353941]

DeMets 1987

DeMets DL. Methods for combining randomized clinical trials: strengths and limitations. *Statistics in Medicine* 1987;**6**(3):341-50. [PMID: 3616287]

DerSimonian 1986

DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177-88. [PMID: 3802833]

Donahoe 2011

Donahoe M. Acute respiratory distress syndrome: a clinical review. *Pulmonary Circulation* 2011;**1**(2):192-211. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 20442122]

Eskesen 2018

Eskesen TG, Baekgaard JS, Steinmetz J, Rasmussen LS. Initial use of supplementary oxygen for trauma patients: a systematic review. *BMJ Open* 2018 July 6;**8**(7):e020880. [DOI: 10.1136/bmjopen-2017-020880.] [PMID: 29982208]

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. [PMID: 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. Art. No: CD009931. [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. Available at gradepro.org.

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. [PMID: 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. [PMID: 26585328]

Hansen 2021

Hansen TE, Christensen RE, Baekgaard J, Steinmetz J, Rasmussen LS. Supplemental oxygen for traumatic brain Injury - a systematic review. Acta anaesthesiologica Scandinavica 2021 December 14. [DOI: 10.1111/aas.14019.] [PMID: 34907522]

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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 11914302]

Higgins 2003

Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60. [PMID: 12958120]

Higgins 2011

Higgins JP, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. *Statistics in Medicine* 2011;**30**(9):903-21. [PMID: 21472757]

Higgins 2016

Higgins JPT, Sterne JAC, Savović J, Page MJ, Hróbjartsson A, Boutron I, Reeves B, Eldridge S. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V (editors). Cochrane Methods. Cochrane Database of Systematic Reviews 2016;(10):Suppl 1. [DOI: dx.doi.org/10.1002/14651858.CD201601]

Higgins 2019

Higgins JPT, Savović J, Page MJ, Sterne JAC. Revised Cochrane risk-of-bias tool for randomized trials (RoB 2). https://www.riskofbias.info/welcome/rob-2-0-tool 22 August 2019.

Higgins 2021

Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021. Available from www.training.cochrane.org/handbook.

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. [PMID: 24881045]

ICH-GCP 1997

International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 25389354]

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. [PMID: 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] [PMID: 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. [PMID: 20530778]

Kallet 2013

Kallet RH, Matthay MA. Hyperoxic acute lung injury. *Respiratory Care* 2013;**58**(1):123-41. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 15470219]

Kulinskaya 2014

Kulinskaya E, Wood E. Trial sequential methods for meta-analysis. *Research Synthesis Methods* 2014;**5**(3):212-20. [PMID: 26052847]

Lefebvre 2021

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021.

Li 2021 (a)

Li L, Zhang Y, Wang P, Chong W, Hai Y, Xu P, Fang F. Conservative versus liberal oxygen therapy for acutely ill medical patients: a systematic review and meta-analysis. *International Journal of Nursing Studies* 2021;**118**:103924. [DOI: 10.1016/j.ijnurstu.2021.103924] [PMID: 33774265]

Li 2021 (b)

Li X, Liub D, Liu C, Mao Z, Liu Y, Yi H, et al. Conservative versus liberal oxygen therapy in relation to all-cause mortality among patients in the intensive care unit: a systematic review of randomized controlled trials with meta-analysis and trial sequential analysis. *Medicina Intensiva* 2021;**16**(261):1-9. [DOI: 10.1016/j.medin.2021.08.006] [PMID: 34526060]

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. [PMID: 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. [PMID: 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] [PMID: 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. [PMID: 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. [PMID: 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).

NCT02321072

NCT02321072. Optimal Oxygenation in the Intensive Care Unit (O2-ICU). clinicaltrials.gov/ct2/show/NCT02321072 (first received 9 December 2014).

NCT02698917

NCT02698917. Carbon Dioxide, Oxygen and Mean Arterial Pressure After Cardiac Arrest and Resuscitation (COMACARE). clinicaltrials.gov/ct2/show/NCT02698917 (first received 4 March 2016).

NCT02713451

NCT02713451. Liberal Oxygenation Versus Conservative Oxygenation in ARDS (LOCO2). clinicaltrials.gov/ct2/show/NCT02713451 (first received 25 February 2016).

NCT02999932

NCT02999932. Pulse Oxygen Saturation (SpO2) Directed Oxygen Therapy (POSDOT). clinicaltrials.gov/ct2/show/NCT02999932 (first received 14 December 2016).

NCT03174002

NCT03174002. Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU). clinicaltrials.gov/ct2/show/NCT03174002 (first recieved 30 May 2017).

NCT03287466

NCT03287466. Targeted OXYgen Therapy in Critical Illness (TOXYC). clinicaltrials.gov/ct2/show/NCT03287466 (first received 20 July 2017).

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. [PMID: 28507176]

Palmer 2019

Palmer E, Post B, lapaukh R, Marra G, MacCallum NS, Brealey D, et al. The association between supraphysiologic arterial oxygen levels and mortality in critically ill patients - a multicenter observational cohort study. *American Journal of Respiratory and Critical Care Medicine* 2019;**200**(11):1373-80. [DOI: 10.1164/rccm.201904-0849OC] [PMID: 31513754]

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] [PMID: 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. [PMID: 23808511]

Patrona 2014

Patrona SD, Zanini A, Aiello M, Adamo D, Casale S, Cherubino F, Raimondi E, Zampogna E, Chetta A, Spanevello A. Estimation of minimum clinically important difference in EQ-VAS score after pulmonary rehabilitation in COPD patients. *European Respiratory Journal* December 23, 2014;**44**(Suppl 58):P3669.

Petersson 2014

Petersson J, Glenny RW. Gas exchange and ventilation-perfusion relationships in the lung. *European Respiratory Journal* 2014;**44**(4):1023-41. [PMID: 25063240]

Pickard 2007

Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility andVAS scores in cancer. *Health and Quality of Life Outcomes* 2007;**5**(70):1-8. [DOI: doi:10.1186/1477-7525-5-70]

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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 30527413]

Review Manager 2020 [Computer program]

The Cochrane Collaboration Review Manager Web (RevMan Web). Version 3.2.1. The Cochrane Collaboration, 2020. Available at revman.cochrane.org.

Rodríguez-Roisin 2005

Rodríguez-Roisin R, Roca J. Mechanisms of hypoxemia. *Intensive Care Medicine* 2005;**31**(8):1017-9. [PMID: 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. [PMID: 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. [PMID: 7760608]

Roussos 2003

Roussos C, Koutsoukou A. Respiratory failure. *European Respiratory Journal* 2003;**47**:3S-14S. [PMID: 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. [PMID: 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. [PMID: 29126260]

Schjørring 2020

Schjørring OL, Jensen AKG, Nielsen CG, Ciubotariu A, Perner A, Wetterslev J, et al. Arterial oxygen tensions in mechanically ventilated ICU patients and mortality: a retrospective, multicentre, observational cohort study. *British Journal of Anaesthesia* 2020 April;**124**(4):420-29. [DOI: 10.1016/j.bja.2019.12.039] [PMID: 32753102]

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. [PMID: 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. [PMID: 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. [PMID: 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] [PMID: 24206183]

STATA 2019 [Computer program]

StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC Stata 16. StataCorp. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC, 2019.

Sterne 2019

Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *British Medical Journal* 2019 August;**366**:l4898. [DOI: 10.1136/bmj.l4898] [PMID: 31462531]

Stub 2015

Stub D, Smith K, Bernard S, Nehme Z, Stephenson M, Bray JE, et al. Air versus oxygen in ST-segment–elevation myocardial infarction. *Circulation* 2015 June;**131**(24):2143-50. [DOI: 10.1161/CIRCULATIONAHA.114.014494] [PMID: 26002889]

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. [PMID: 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] [PMID: 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. [PMID: 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. [PMID: 22923432]

TSA 2011 [Computer program]

Copenhagen Trial Unit TSA - Trial Sequential Analysis. Version 0.9 Beta. Copenhagen: Copenhagen Trial Unit, 2011. www.ctu.dk/tsa/downloads.aspx.

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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. Art. No: CD008884. [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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 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. [PMID: 22404056]

Young 2017

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 pilot phase. *Critical Care and Resuscitation* 2017 December;**19**(4):344-54. [PMID: 29202261]

Young 2021

Young PJ, Bailey M, Bellomo R, Bernard S, Bray J, Jakkula P, et al. Conservative or liberal oxygen therapy in adults after cardiac arrest: An individual-level patient data meta-analysis of randomised controlled trials. *Resuscitation* 2020 ;**157**:15-22. [DOI: 10.1016/j.resuscitation.2020.09.036.] [PMID: 33058991]

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. [PMID: 17349918]

Zanini 2015

Zanini A, Aiello M, Adamo D, Casale S, Cherubino F, Patrona SD, et al. Estimation of minimal clinically important differencein EQ-5D visual analog scale score after pulmonary rehabilitation in subjects with COPD. *Respiratory Care* January 2015;**60**(1):88-95. [DOI: 10.4187/respcare.03272] [PMID: 25336531]

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. [PMID: 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 (protocol). *Cochrane Database of Systematic Reviews* 2017 April, Issue 4. Art. No: CD012631. [DOI: 10.1002/14651858.CD012631]

Barbateskovic 2019

Barbateskovic M, Schjørring OL, Krauss SR, Jakobsen JC, Meyhoff CS, Dahl RM et al. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit (review). *Cochrane Database of Systematic Reviews* 2019 November 27, Issue 11. Art. No: CD012631. [DOI: 10.1002/14651858.CD012631.pub2] [PMID: 31773728]

* Indicates the major publication for the study

ISSN (online): 2246-1302
ISBN (online): 978-87-7573-948-6

AALBORG UNIVERSITY PRESS