Aalborg Universitet

Aalborg University
DENMARK

Long-term effects of lower versus higher oxygenation levels in adult ICU patients—A systematic review

Crescioli, Elena; Krejberg, Kirsten Uldal; Klitgaard, Thomas Lass; Nielsen, Frederik Mølgaard; Barbateskovic, Marija; Skrubbeltrang, Conni; Møller, Morten Hylander; Schjørring, Olav Lilleholt; Rasmussen, Bodil Steen

[Link to publication from Aalborg University](#)

REVIEW ARTICLE

acta Anaesthesiologica Scandinavica

Long-term effects of lower versus higher oxygenation levels in adult ICU patients—A systematic review

Elena Crescioli[1,2] | Kirsten Uldal Krejberg[2] | Thomas Lass Klitgaard[1,2] |
Frederik Mølgaard Nielsen[1,2] | Marija Barbateskovic[3] | Conni Skrubbeltrang[4] |
Morten Hylander Møller[5] | Olav Lilleholt Schjørring[1,2] | Bodil Steen Rasmussen[1,2]

[1]Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Aalborg, Denmark

[2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

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

[4]Medical Library, Aalborg University Hospital, Aalborg, Denmark

[5]Department of Intensive Care, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark

Correspondence
Bodil Steen Rasmussen, Department of Anaesthesia and Intensive Care, Aalborg University Hospital, Hobrovej 18-22, 9100 Aalborg, Denmark.
Email: bodil.steen.rasmussen@rn.dk

Funding information
Danish Ministry of Higher Education and Science; Novo Nordisk Fonden

Abstract

Background: Oxygen therapy is a common treatment in the intensive care unit (ICU) with both potentially desirable and undesirable long-term effects. This systematic review aimed to assess the long-term outcomes of lower versus higher oxygenation strategies in adult ICU survivors.

Methods: We included randomised clinical trials (RCTs) comparing lower versus higher oxygen supplementation or oxygenation strategies in adults admitted to the ICU. We searched major electronic databases and trial registers. We included all non-mortality long-term outcomes. Prespecified co-primary outcomes were the long-term cognitive function measures, the overall score of any valid health-related quality of life (HRQoL) evaluation, standardised 6-min walk test, and lung diffusion capacity. The protocol was published and prospectively registered in the PROSPERO database (CRD42021223630).

Results: The review included 17 RCTs comprising 6592 patients, and six trials with 825 randomised patients reported one or more outcomes of interest. We observed no difference in cognitive evaluation via Telephone Interview for Cognitive Status (one trial, 409 patients) (mean score: 30.6 ± 4.5 in the lower oxygenation group vs. 30.4 ± 4.3 in the higher oxygenation group). The trial was judged at overall high risk of bias and the certainty of evidence was very low. Any difference was neither observed in HRQoL measured via EuroQol 5 dimensions 5 level questionnaire and EQ Visual Analogue Score (one trial, 499 patients) (mean score: 70.1 ± 22 in the lower oxygenation group vs. 67.6 ± 22.4 in the higher oxygenation group). The trial was judged as having high risk of bias, the certainty of evidence was very low. No trial reported neither the standardised 6-min walk test nor lung diffusion test.

Conclusion: The evidence is very uncertain about the effect of a lower versus a higher oxygenation strategy on both the cognitive function and HRQoL. A lower versus a higher oxygenation strategy may have a little to no effect on both outcomes but the certainty of evidence is very low. No evidence was found for the effects on the standardised 6-min walking test and diffusion capacity test.

Editorial Comment

This trustworthy systematic review assessed randommized clinical trials focused on oxygenation treatment goals in critically ill study participants. Several important clinical outcomes were included in the analysis, but the strength of the evidence was found to be low in the available trials, limiting the ability to draw conclusions on relative benefit for higher or lower oxygenation treatment goals in this context.

1 | INTRODUCTION

Along with the increasing number of survivors after admission to the intensive care unit (ICU) the focus of scientific research has expanded to also include post-ICU long-term outcomes.[1] ICU survivors can be affected by longstanding organ dysfunctions after hospital discharge.[2] Psychiatric complications and cognitive impairment occur frequently, and up to half of ICU patients discharged from hospital have altered mental status.[2] Pulmonary dysfunctions have mostly been studied in survivors of acute respiratory distress syndrome (ARDS) with results showing predominantly mild impairments.[3,4] ICU-acquired weakness, a neuromuscular complication of critical illness associated with longer durations of mechanical ventilation and in-hospital length of stay, is also acknowledged.[5] Finally, limitations in physical function, typically measured by surveying patients in the activities of daily living, are commonly reported and may be irreversible.[2]

Among the different interventions undertaken during ICU stay, oxygen is the most commonly prescribed drug.[6] Supplemental oxygen is given to either prevent or treat hypoxaemia, and hyperoxaemia with supra-normal oxygen levels has often been tolerated as a safety buffer.[7] The growing interest in targeted supplemental oxygen therapy arises from concerns about side-effects from hyperoxia and fears of hypoxaemia. In the last decade, several large-scale randomised clinical trials (RCTs) have investigated lower versus higher oxygenation strategies in the ICU setting, reporting conflicting results regarding mortality.[8–13] However, evidence on the long-term effects of oxygen therapy in adult ICU survivors is lacking.

Presently, no systematic reviews investigating the impact of oxygen therapy in the ICU, have reported any other long-term outcomes than mortality and quality of life.[14] Therefore, we conducted a systematic review with meta-analyses on oxygen supplementation's potential impact on all long-term non-mortality outcomes. The primary objective of this review was to assess the long-term effects of lower versus higher oxygen supplementation or oxygenation levels in adult ICU survivors. We a priori hypothesised that lower oxygenation strategies would result in poorer long-term cognitive function, whereas higher oxygenation strategies would result in poorer long-term pulmonary function, poorer

standardised 6-min walk test,[15] and reduced health-related quality of life (HRQoL).[16]

2 | METHODS

This systematic review was conducted according to the prespecified and published protocol.[16] We prospectively registered the protocol in the international prospective register of systematic reviews database (PROSPERO) (CRD42021223630), used the methodology of the Cochrane Handbook,[17] and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement.[18] The PRISMA checklist is available in the Supporting Information Materials S1.

2.1 | Eligibility criteria

RCTs comparing a lower and a higher oxygenation strategy were included. Oxygenation strategies were defined by fraction of inspired oxygen (FiO_2), including separate oxygen flow levels in open systems, or by oxygenation targets or levels measured by arterial partial pressure of oxygen (PaO_2), or arterial or peripheral oxygen saturation (SaO_2 or SpO_2). We included adult patients and only if randomised upon or after ICU admission. To ensure inclusion of all relevant trials, no thresholds for oxygenation for the two groups were determined beforehand. Quasi-randomised trials, individual patient cross-over trials, and trials on hyperbaric oxygen or hypoxaemia were excluded.

2.2 | Search methods

We based our search strategy on the review by Barbateskovic et al.[14] The following databases were searched: Cochrane Central Register of Controlled Trials, MEDLINE, Embase, Science Citation Index, BIOSIS Previews, Latin American, and Caribbean Health Science Information database. The search was updated on 6 January 2022. Detailed search strategies are listed in the Supporting Information S1. In addition, reference lists of relevant reviews and papers were manually screened, and we also searched trial registers.[16]

2.3 | Trial selection and data extraction

Four authors (T.L.K., F.M.N., M.B., and O.L.S.) independently and in pairs screened titles and abstracts. Reports deemed potentially relevant were obtained in full text and assessed for inclusion. Two authors (K.U.K. and E.C.) independently extracted predefined data from the included trials using a standardised data collection form (Supporting Information S1). Any disagreement was resolved by consensus, or upon the involvement of a co-author (M.B., O.L.S., or B.S.R.).

2.4 | Risk of bias assessment

Two authors (E.C. and K.U.K.) independently assessed risk of bias according to the Cochrane Handbook for Systematic Reviews of Interventions using the revised Risk of Bias tool 2.[19,20] This was done for each trial reporting at least one outcome of interest, and for each individual outcome. We assessed risk of bias in all the five mandatory domains: bias arising from the randomisation process; bias due to deviations from intended interventions; bias due to missing outcome data; bias in measurement of the outcome; and bias in the selection of the reported results. When assessing the domain 'bias due to deviations from intended interventions', we judged the effect of assignment to the intervention (i.e., intention to treat effect). Each domain was adjudicated as 'low risk of bias', 'some concerns', or 'high risk of bias'. Further details on risk of bias classification are explained in Supporting Information S1.

2.5 | Data synthesis

We calculated the mean difference with a 95% confidence interval (CI) for continuous data. For similar outcomes measured by different scales we calculated standardised mean difference presented with 95% CIs. We calculated risk ratios (RR) with 95% CIs for dichotomous outcomes.

2.6 | Assessment of heterogeneity

Heterogeneity was assessed by visual inspection of forest plots, statistical heterogeneity was assessed using χ^2 test with significance at a p value below 10% and the quantities of heterogeneity was measured by calculations of I^2, where a I^2 >50% was considered substantial.[21] The tool 'Clinical Diversity in Meta-analyses' (CDIM) was used to assess the clinical diversity of each meta-analysis within the following four domains: setting, population, intervention, and outcome diversity.[22]

2.7 | Outcomes

As prespecified, all long-term outcomes excluding mortality were included.[23] 'Long-term' was defined as any time point after hospital discharge. The predefined co-primary outcomes were: any cognitive function measure, the overall score of any valid HRQoL assessment, the standardised 6-min walk test,[15] and lung diffusion capacity test.[24] Any additional long-term outcomes were reported as exploratory. For all outcomes, trial results reported at longest follow-up were used in the analyses.

2.8 | Meta-analysis

If two or more RCTs with comparable outcome measures were included, we assessed intervention effects with both random-effects[25–27] and fixed-effect meta-analysis.[28] We used the more conservative point estimate (being the one closest to the null-effect) of the two with the highest p value. Analyses were conducted using *STATA statistical software* (Stata Nordic, version 17), and results are illustrated using forest plots. As prespecified,[16] we performed adjustment for multiplicity, considering statistically significant a p value below 2%, equivalent to an adjusted CI of 98%.[29] Exploratory outcomes were not adjusted for multiplicity, and a $p < 5\%$ was assumed significant. In accordance with our published protocol,[23] when analysing our co-primary outcomes, we also planned to perform a subgroup analysis according to the type of ICU population (i.e., medical vs. surgical vs. mixed) and a sensitivity analysis (i.e., best–worst and worst–best case scenarios).[16]

2.9 | Trial sequential analysis

We analysed our prespecified co-primary outcomes with trial sequential analysis (TSA).[30] For dichotomous outcomes, we estimated the required information size (RIS) based on the observed proportion of patients with an outcome in the control group (the cumulative proportion of patients with an event in the control groups relative to all patients in the control groups), a RR reduction or a RR increase of 20%, an α of 2% for all our outcomes, a β of 10% (i.e., power of 90%), and the observed diversity as suggested by the trials in the meta-analysis. For continuous outcomes, we used the observed standard deviation (SD) in the control group, the observed SD/2 as tested difference, an α of 2% for all outcomes, a β of 10%, and the observed diversity as suggested by the trials in the meta-analysis. When analysing EQ-VAS, we performed an additional post hoc TSA based on a proposed minimum important difference for the EQ-VAS of 7.[31–33]

2.10 | GRADE assessment

We assessed the certainty of evidence for all outcomes according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.[34–36] We present the results of the GRADE assessment in Table 1. We appraised the certainty of the evidence and our confidence in the effect estimates based on risk of bias, imprecision, inconsistency, indirectness, and publication bias. The

TABLE 1 Summary of findings

Certainty assessment							No. of patients		Effect		Certainty
No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lower	Higher	Relative (95% CI)	Absolute (95% CI)	
Cognitive function[a]											
1[11]	RCT	Serious[b]	Serious[c]	Serious[d]	Not serious	None	203	206		Mean score: 30.6 ± 4.5 in the lower oxygenation group vs. 30.4 ± 4.3 in the higher oxygenation group	⊕○○○ very low
Health-related quality of life[e]											
1[11]	RCT	Serious[b]	Serious[c]	Serious[d]	Not serious	None	246	253		Mean score: 70.1 ± 22 in the lower oxygenation group vs. 67.6 ± 22.4 in the higher oxygenation group	⊕○○○ very low
Standardised 6-min walk test											
0	Not available	Not available	Not available	Not available	Not available	Not available	Not available	Not available	Not estimable	Not available	Not estimable
Diffusion capacity test											
0	Not available	Not available	Not available	Not available	Not available	Not available	Not available	Not available	Not estimable	Not available	Not estimable
Pulmonary function											
1[37]	RCT	Serious[b]	Serious[c]	Serious[d]	Serious[f]	None	12	12		Mean score of FEV1: 0.56 ± 0.18 in the lower oxygen group vs. 0.5 ± 0.32 in the higher oxygenation group. Mean score of FVC: 1.6 ± 0.86 in the lower oxygenation group vs. 1.05 ± 0.59 in the higher oxygenation group	⊕○○○ very low
Functional outcome assessed using Barthel Index[g]											
2[38,39]	RCT	Serious[b]	Not serious	Serious[h]	Serious[f]	None	58	58		MD −8.50 (−14.99 to −2)	⊕○○○ very low
Functional outcome assessed using modified Rankin Scale[i]											
2[38,39]	RCT	Serious[b]	Not serious	Serious[h]	Serious[f]	None	59	60		MD 0.83 (0.32–1.35)	⊕○○○ very low
Functional outcome assessed using Glasgow Outcome Scale or extended Glasgow Outcome Scale[j]											
3[11,38,39]	RCT	Serious[b]	Not serious	Serious[h]	Serious[k]	None	118/226 (52.2%)	112/228 (49.1%)	RR 0.95 (0.81–1.12)	5 fewer per 1000 (from 19 fewer to 12 more)	⊕○○○ very low
Functional outcome assessed using Cerebral Performance Category[j]											
1[40]	RCT	Not serious	Serious[c]	Serious[d]	Serious[f]	None	42/61 (68.9%)	36/59 (61.0%)	Not estimable	42/61 (68.9%) in the lower oxygenation group, and 36/59 (61.0%) in the higher oxygenation group	⊕○○○ very low

(Continues)

TABLE 1 (Continued)

Certainty assessment							No. of patients		Effect			
No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lower	Higher	Relative (95% CI)	Absolute (95% CI)		Certainty
Return to work[m]												
1[11]	RCT	Not serious	Serious[c]	Serious[d]	Serious[f]	None	112	108	Not estimable	77/112 (68.8%) in the lower oxygenation group and 66/108 (61.1%) in the higher oxygenation group		⊕◯◯◯ very low

Note: 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.

Abbreviations: ADL, activities of daily living; CI, confidence interval; MD, mean difference; RR, risk ratio.

[a]Measured by the Telephone Interview for Cognitive Status, which is an 11-item global mental status test with higher scores meaning a better performance and a maximum score of 41.[41]

[b]Trial was judged at overall high risk of bias.

[c]We cannot reject inconsistency due to inclusion of only one trial.

[d]We cannot reject indirectness due to inclusion of only one trial.

[e]Measured by the the EuroQol five dimension five level (EQ-5D-5L) questionnaire, including the Visual Analogue Scale (EQ-VAS). EQ-VAS asks the patients to self-rate their perceived overall health on a scale from 0 (i.e., 'the worst health you can imagine') to 100 (i.e., 'the best health you can imagine').[42,43]

[f]Few patients included in the trial.

[g]The Barthel Index is a 10-item scale of basic ADL. The 10 items focus on self-care and mobility, and the individual scores of the 10 items are summed to a maximum possible score of 100 (independent) and a minimum of 0 (totally dependent).[44]

[h]Differences in inclusion criteria and inspiratory oxygen fraction in both the experimental and control group between trials.

[i]The modified Rankin Scale is a global disability scale to measure the level of functional independence in daily activities. It consists of seven grades going from 0 (i.e., no symptoms) to 6 (i.e., death).[45]

[j]The Glasgow Outcome Scale (GOS) is a 5-point scale going from 1 (i.e., death) to 5 (i.e., low disability) (ref.). The extended GOS (eGOS) is a 8-point scale, from 1 being death to 8 meaning upper good recovery (i.e., full recovery).[46,47] Data and meta-analysis are based on a dichotomised scale (i.e., good vs. poor good outcome), defining a good outcome a GOS ≥4 and eGOS ≥5. Data are reported as the proportion of patients with a poor outcome.

[k]Serious imprecision: 95% CI of RR indicated appreciable and non-appreciable benefit for lower oxygenation strategy.

[l]The cerebral performance category (CPC) is a 5-point scale that ranges from 1 (i.e., good cerebral performance) to 5 (i.e., brain death).[48] Data are reported as the proportion of patients with a good outcome defined as a CPC ≤2 by trialists.

[m]Return to work defined as the employment status among survivors with a paid employment at randomisation.
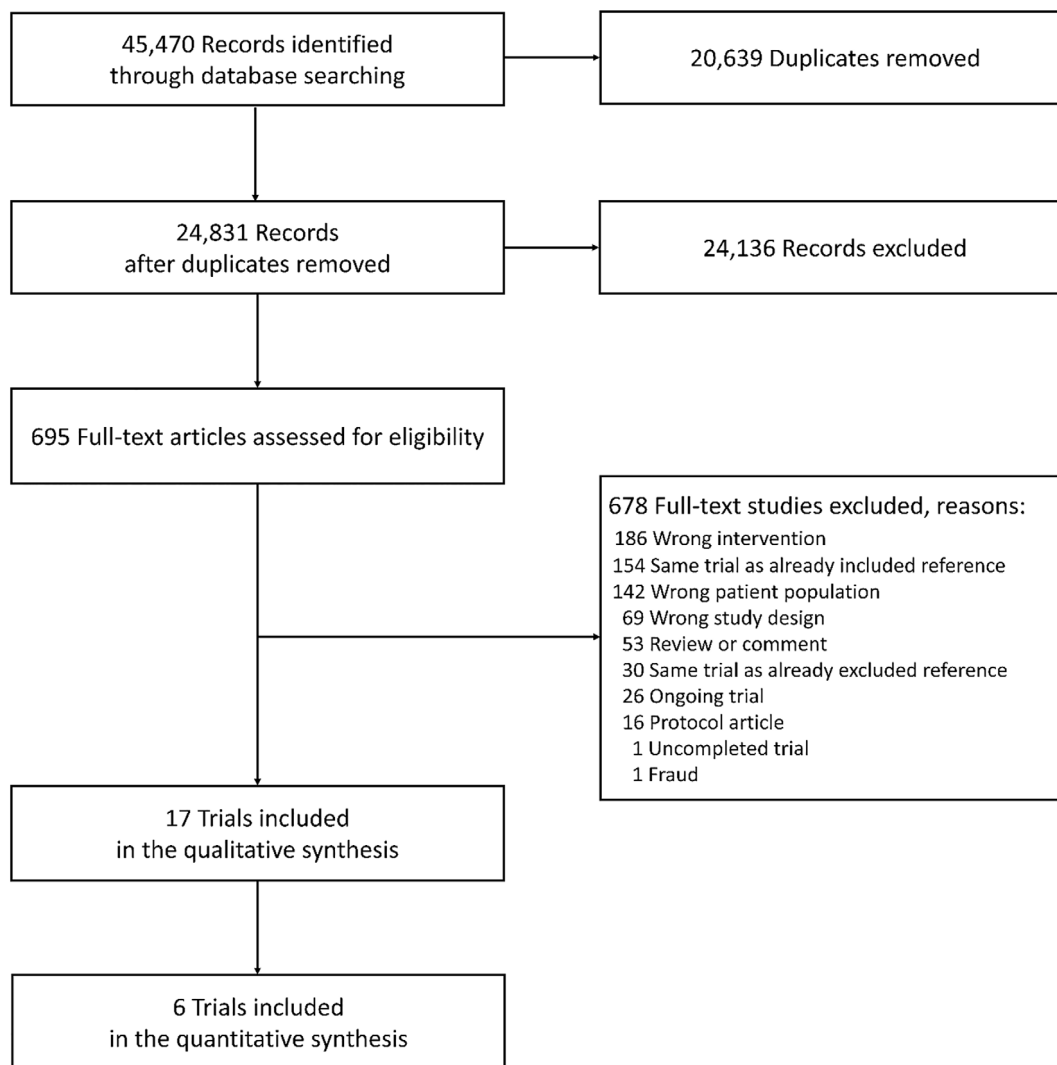
```
┌─────────────────────────────────┐        ┌─────────────────────────────────┐
│   45,470 Records identified     │───────▶│   20,639 Duplicates removed     │
│   through database searching    │        │                                 │
└─────────────────────────────────┘        └─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐        ┌─────────────────────────────────┐
│      24,831 Records             │───────▶│   24,136 Records excluded       │
│   after duplicates removed      │        │                                 │
└─────────────────────────────────┘        └─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│  695 Full-text articles assessed │
│         for eligibility          │
└─────────────────────────────────┘
                 │                          ┌─────────────────────────────────────────────┐
                 │                          │ 678 Full-text studies excluded, reasons:    │
                 │                          │  186 Wrong intervention                     │
                 │                          │  154 Same trial as already included reference│
                 │─────────────────────────▶│  142 Wrong patient population               │
                 │                          │   69 Wrong study design                     │
                 │                          │   53 Review or comment                       │
                 │                          │   30 Same trial as already excluded reference│
                 │                          │   26 Ongoing trial                          │
                 │                          │   16 Protocol article                       │
                 │                          │    1 Uncompleted trial                      │
                 │                          │    1 Fraud                                  │
                 ▼                          └─────────────────────────────────────────────┘
┌─────────────────────────────────┐
│     17 Trials included          │
│   in the qualitative synthesis  │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│      6 Trials included          │
│   in the quantitative synthesis │
└─────────────────────────────────┘
```

FIGURE 1   Preferred Reporting Items for Systematic Reviews and Meta-Analysis flowchart[18]

overall certainty of evidence was graded as 'very low', 'low', 'moderate', or 'high'.

3 | RESULTS

3.1 | Results of the search and selection of trials

We identified 45,470 titles and assessed 695 full texts for eligibility. Ultimately, 17 RCTs comprising 6592 patients were included (Figure 1).

3.2 | Characteristics of included trials

Eleven RCTs did not report any outcome of interest[8–10,12,13,49–54]; six trials were included in the quantitative analysis.[11,37–40,55] The number of participants ranged from 34 to 2928, and all RCTs included adults admitted to the ICU: nine trials included multidisciplinary ICU patients[8–13,51,52]; four trials included only medical ICU patients[37,38,40,54];

and two trials included only surgical ICU patients.[49,55] Two trials did not report the type of ICU to which patients were admitted.[39,50] The trials were conducted across different countries in Europe, Asia, Australia, and New Zealand. All trials assessed lower versus higher oxygenation strategies using either FiO_2 or arterial oxygenation targets, or a combination hereof. The definitions of lower versus higher oxygenation strategies differed widely between the trials. In the lower oxygenation group, FiO_2 ranged from 21% to 50%, whereas in the higher oxygenation group FiO_2 ranged from 30% to 100%. Duration of the intervention also differed, ranging from 1 h to 90 days. Further details are presented in Table 2 and Supporting Information S1.

3.3 | Risk of bias

The trial that reported both cognitive function and HRQoL was judged as being at overall high risk of bias.[11] The trial that reported pulmonary function was also deemed at overall high risk of bias.[37] All but one trial[40] that reported on functional outcomes were evaluated at

TABLE 2 Characteristics of included trials

Trial	Country	Setting	Sample size	Duration	Lower oxygenation group FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	Higher oxygenation group FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	Maximum follow-up
Asfar et al.[9]	France	Adults with septic shock admitted to multidisciplinary ICU	442	24 h			88%–95%	1.00			6 months
Barrot et al.[10]	France	Adults with ARDS admitted to mixed disciplinary ICUs	205	7 days or until extubation (before 7 days)		7.3–9.3 kPa	88%–92%		12–14 kPa	≥96%	90 days
Gelissen et al.[13]	The Netherlands	Adults with systemic inflammation admitted to multidisciplinary ICU	574	14 days or until ICU discharge or death (before 14 days)		8–12 kPa			14–18 kPa		90 days
Girardis et al.[8]	Italy	Adults admitted to multidisciplinary ICU	480	ICU stay (median, 144)		9.3–13.3 kPa	94%–98%		≤20 kPa	97%–100%	60 days
Gomersall et al.[37]	Hong Kong	Adults with AECOPD admitted to multidisciplinary ICU	36	Length of hospital stay		>6.6 kPa			>9.0 kPa		Not specified
Ishii et al.[49]	Japan	Mechanically ventilated patients admitted to surgical ICU	44	1 h	Expected FiO$_2$ to achieve a PaO$_2$ of 13.3 kPa	13.3 kPa		100%			Day 5 postextubation
Jakkula et al.[40]	Finland and Denmark	Mechanically ventilated adults admitted to the ICU after OHCA	123	36 h		10–15 kPa	95%–98%		20–25 kPa		6 months
Jun et al.[50]	Not specified	Not specified	87	96 h	30%–50%			50%–70% within 48 h, afterwards FiO$_2$ was gradually decreased from 50% to 40% over 48 h			14 days
Lång et al.[55]	Finland	Mechanically ventilated adults with TBI admitted to the ICU	65	Throughout mechanical ventilation for a maximum of 14 days	40%			70%			6 months

TABLE 2 (Continued)

Trial	Country	Setting	Sample size	Interventions		Lower oxygenation group			Higher oxygenation group			Maximum follow-up
				Duration		FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	
Mackle et al.[11]	Australia and New Zealand	Mechanically ventilated adults admitted to multidisciplinary ICUs	1000	Until death or discharge from the ICU, or Day 28 post-randomisation				90%–96%	≥30%		90%–100%	180 days
Martin et al.[51]	England	Mechanically ventilated adults admitted to multidisciplinary ICU	34	Until extubation, formation of tracheostomy, transfer to another ICU or death				88%–92%			≥96%	90 days
Mazdeh et al.[38]	Iran	Adults with stroke initially referred to the Department of Neurology, but admitted to the ICU	51	12 h		No supplemental oxygen			50%			6 months
Panwar et al.[52]	Australia, New Zealand, and France	Mechanically ventilated adults admitted to a multidisciplinary ICU	104	Entire duration of mechanical ventilation (median, 114 h)				88%–92%			≥96%	90 days
Schjørring et al.[12]	Denmark, Switzerland, Finland, the Netherlands, Norway, the United Kingdom, and Iceland	Adults with acute hypoxaemic respiratory failure at multidisciplinary ICUs	2928	Until discharge from ICU, or death (including ICU readmission). Maximum 90 days			8 kPa			12 kPa		90 days
Taher et al.[39]	Iran	Adults with TBI initially referred to the emergency department, but who were admitted to the ICU	68	6 h		50%			80%			6 months

(Continues)

TABLE 2 (Continued)

Trial	Country	Setting	Sample size	Interventions							Maximum follow-up
				Lower oxygenation group				Higher oxygenation group			
				Duration	FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	FiO$_2$/O$_2$ flow	PaO$_2$	SaO$_2$/SpO$_2$	
Yang et al.[53]	China	Multidisciplinary ICU	214	Not specified		As low as possible	90%–95%		≥30%	96%–100%	28 days
Yang et al.[54]	China	Mechanically ventilated adults with severe pneumonia admitted to the ICU	106	Not specified		9.3–13.3 kPa	90%–92%		>20 kPa	>96%	In-hospital

Abbreviations: AECOPD, acute exacerbation of chronic obstructive pulmonary disease; ARDS, acute respiratory distress syndrome; FiO$_2$, fraction of inspired oxygen; ICU, intensive care unit; O$_2$, oxygen; OHCA, out-of-hospital cardiac arrest; PaO$_2$, arterial partial pressure of oxygen; SpO2, peripheral oxygen saturation; TBI, traumatic brain injury.

overall high risk of bias.[11,38,39,55] The trial that reported return to work was judged as having overall low risk of bias.[11] The domains mostly affected by risk of bias were bias due to deviations from intended interventions, bias due to missing outcome data, and bias due to selection of the reported results (Tables S1–S9).

3.4 | Effect of interventions

The outcome definitions from contributing trials can be found in the Supporting Information S1.

3.4.1 | Co-primary outcomes

No trials included in this review reported the standardised 6-min walking test or the diffusion capacity test (Table 1).

Cognitive function

Only one trial ($n = 409$ patients) reported evaluation of cognitive function at 180 days, using the Telephone Interview for Cognitive Status (TICS).[11] The trial was judged at overall high risk of bias and no difference in TICS scores was found between the intervention groups (Table 1).[11] TSA of TICS demonstrated that with an anticipated mean difference of 2.2 points the population size in the identified trial exceeded the RIS, thus no TSA graph was produced. The certainty of evidence was very low (Table 1).

Quality of life

One trial reported HRQoL at 180 days ($n = 499$ patients), using the EuroQol five dimensions five levels questionnaire including the Visual Analogue Scale (EQ-VAS). The trial was judged at an overall high risk of bias, and no difference in EQ-VAS scores was found between the lower and higher oxygenation groups (Table 1).[11] TSA of HRQoL demonstrated that with an anticipated mean difference of 10.6 points the population size in the identified trial exceeded the RIS, thus no TSA graph was produced. In a post hoc TSA with a mean difference of 7 points, the population size reached 95.2% of the RIS (Figure S1). The certainty of evidence was very low (Table 1).

3.4.2 | Exploratory outcomes

Pulmonary function

One trial ($n = 24$) reported pulmonary function tests, being the FEV$_1$ and the FVC. For both outcome measures, the trial was judged at overall high risk of bias. No differences between the trial groups were found, and the time point of the follow-up was not specified.[37] The certainty of evidence was very low (Table 1).

Functional outcomes

Barthel Index. The Barthel Index was reported by two trials ($n = 116$ patients), both judged at overall high risk of bias.[38,39] Meta-analysis of
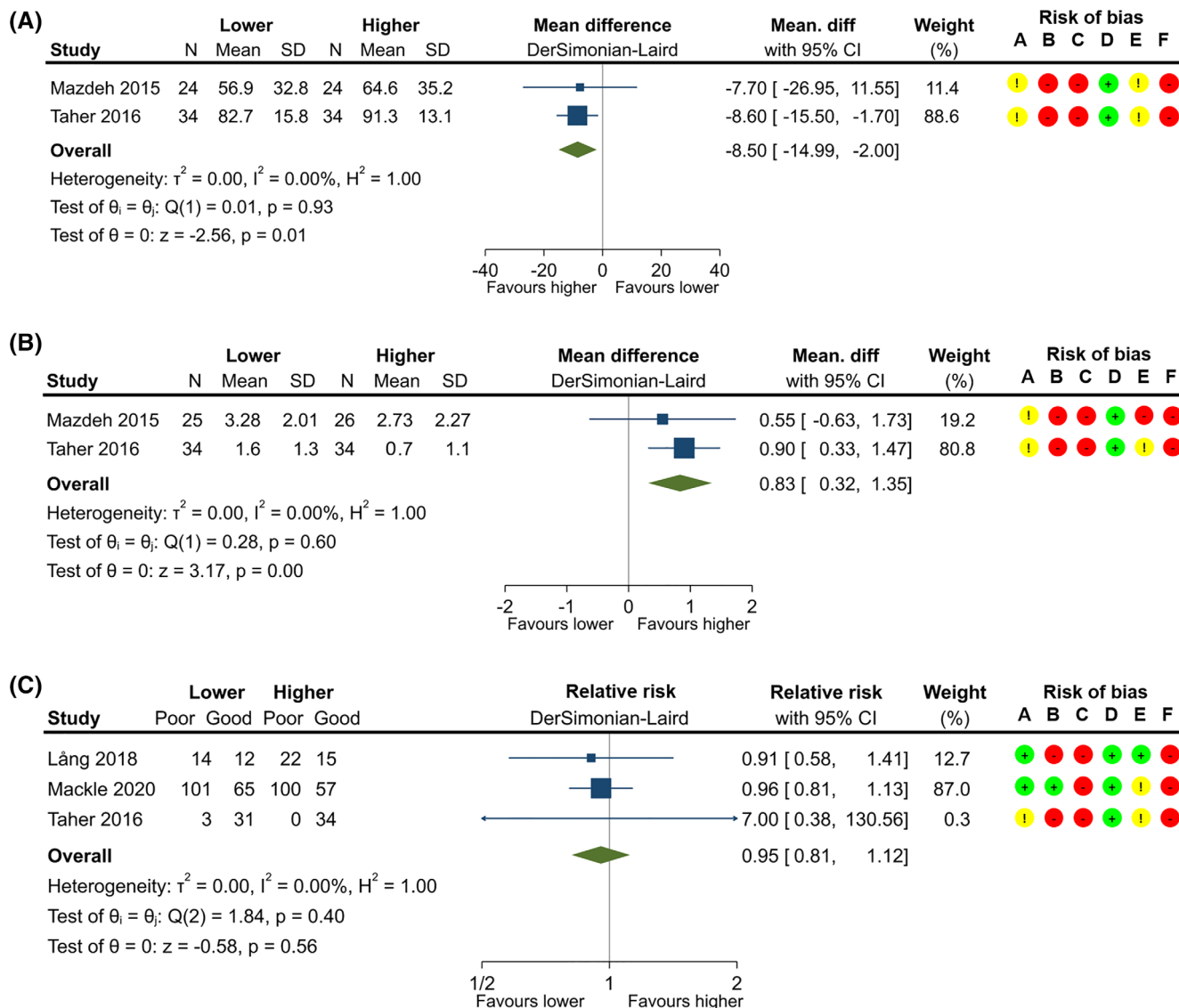
FIGURE 2 Forest plots on long-term outcomes measured by Barthel Index (A), modified Rankin Scale (B), and Glasgow Outcome Scale (GOS) or extended GOS (C), respectively. Risk of bias legend. (A) Bias arising from the randomisation process. (B) Bias due to deviations from intended interventions. (C) Bias due to missing outcome data. (D) Bias in measurement of the outcome. (E) Bias in selection of the reported result. (F) Overall bias. Size of squares for risk ratio reflects weight of trial in pooled analysis. Horizontal bars represent 95% CI. CI, confidence interval; RR, risk ratio

these trials showed a favourable outcome in the higher oxygenation group (mean difference: −8.50, 95% CI: −14.99 to −2, I^2 = 0%) (Figure 2A). CDIM showed a moderate clinical diversity (Table S10). The certainty of evidence was very low (Table 1).

Modified Rankin Scale. The modified Rankin Scale (mRS) was reported by two trials (n = 119 patients), both at overall high risk of bias.[38,39] Meta-analysis showed a significantly better outcome in the higher oxygenation group (mean difference: 0.83, 95% CI: 0.32–1.35, I^2 = 0%) (Figure 2B). CDIM showed a moderate clinical diversity (Table S11). The certainty of evidence was very low (Table 1).

Glasgow Outcome Scale and extended Glasgow Outcome Scale. Glasgow Outcome Scale (GOS) was reported by one trial which was deemed at

overall high risk of bias for this outcome,[39] while extended Glasgow Outcome Scale (eGOS) was reported by two trials, both deemed at high risk of bias (n = 454 patients).[11,55] We conducted a meta-analysis of the three trials dichotomising both scales (i.e., good vs. poor outcome), defining a good outcome as GOS ≥4 and eGOS ≥5. Raw data are presented in Table S12. The meta-analysis showed no significant difference between the lower and higher oxygenation groups (RR: 0.95, 95% CI: 0.81–1.12, I^2 = 0%) (Figure 2C). CDIM showed a moderate clinical diversity (Table S13). The certainty of evidence was very low (Table 1).

Cerebral performance category

One trial (n = 120 patients) judged at overall low risk of bias, reported cerebral performance category (CPC) without demonstrating any

difference between the trial groups (Table 1).[40] The certainty of evidence was very low (Table 1).

Return to work

One trial (*n* = 220 patients) judged at overall low risk of bias, reported return to work at 180 days within the population of survivors with paid employment at randomisation. No difference was found between the lower versus higher oxygenation groups (Table 1).[11] The certainty of evidence was very low (Table 1).

Subgroup and sensitivity analyses

Due to very limited data, we were unable to conduct any of the pre-planned subgroup and sensitivity analyses.

4 | DISCUSSION

In this systematic review of RCTs reporting long-term effects of lower versus higher oxygenation levels in adult ICU patients, the quantity and quality of evidence was very low with no firm evidence for benefit or harm. Six trials reported one or more outcomes of interest with a total of 825 randomised participants contributing with data. Only two prespecified co-primary outcomes, long-term cognitive function evaluation and HRQoL, were reported on, and by only one trial.[11] Thus, no meta-analysis for any co-primary outcomes could be performed.

In a Cochrane review on oxygenation strategies in adult ICU populations,[14] a higher as compared to a lower oxygenation strategy was found to possibly increase mortality and the incidence of SAEs. However, findings were based on very low certainty of evidence. None of the included trials reported results on HRQoL assessments.[14]

Our findings further highlight the lack of sufficient evidence concerning the long-term effects of different oxygenation strategies in adult ICU survivors. The existing knowledge is thus unable to inform current clinical practice. Presently, long-term outcomes in ICU survivors have mostly been investigated in observational studies with hypoxaemia during ICU stay being highlighted as a key contributor to cognitive dysfunction in ARDS survivors.[56,57] Due to an increasing awareness of morbidity among ICU survivors, the investigation of long-term outcomes in clinical trials conducted in the intensive care setting has, in the last decades, been advocated as an urgent matter.[58] Although the measurement of long-term outcomes can be difficult, mortality and other short-term outcomes cannot be assumed to be their comparable proxies. In our case, several large-scale RCTs have explored differences between lower and higher oxygenation strategies in terms of mortality,[8,10–13] but only one trial also reported long-term outcomes in the population of survivors.[11] The few other RCTs, exploring long-term outcomes, are hampered by small trial populations,[37–40,55] and serious issues of bias[11,37–39,55]; the most frequent being missing data due to patients lost to follow-up.

A challenge for future research is to design RCTs which systematically incorporate longer follow-up (e.g., 6 months) combined with standardised outcomes allowing for between-trial comparisons. Loss

to follow-up and trial withdrawal may undermine the statistical power of clinical trials and it will become crucial to understand such loss to assess the full spectrum of disability among ICU survivors.

The current review holds several strengths. It is reported in accordance with the PRISMA statement,[18] and the methodology is based on the Cochrane Handbook for Systematic Reviews of Intervention,[17,19] and the GRADE approach.[34] Moreover, we used a predefined and rigorous search strategy,[14] identifying all relevant trials, and contacting the trial investigators if additional information was needed. To enhance transparency, the protocol was published in advance,[16] and was prospectively registered in the PROSPERO database.

The review also holds limitations, the primary being that we, due to the lack of international consensus regarding targeted oxygen therapy, did not a priori define oxygenation thresholds. To define the interventions, FiO_2, PaO_2, and SpO_2/SaO_2 could be used separately, or in combination in the same trial, generating a significant heterogeneity regarding the applied interventions. Moreover, the trials included in the review vastly differed in several other domains, for example, setting, ICU population, and timing and duration of the intervention. Although the statistical heterogeneity seemed low, the clinical diversity was evaluated as moderate using the CDIM tool,[22] but we were unable to perform subgroup analyses due to limited data. As expected, we also found an extreme diversity in outcome measures, which are difficulty to mutually compare, due to the populations they are addressing, and distinct scoring systems. Moreover, when using ordinal scales such as mRS, GOS, and CPC, a dichotomisation (i.e., favourable vs. unfavourable outcome) of the measure has often been used for analyses in trials and meta-analyses,[59] potentially affecting the statistical power of the results.[59] Finally, it is important to mention that we did not include mortality at the longest follow-up in this review, since it has been explored previously,[14] and our focus was on ICU survivors.

5 | CONCLUSION

The evidence is very uncertain about the effect of a lower versus a higher oxygenation strategy on both the cognitive function and HRQoL. A lower versus a higher oxygenation strategy may have a little to no effect on both outcomes but the certainty of evidence is very low. No evidence was found for the effects on the standardised 6-min walking test and diffusion capacity test.

on previous versions of the manuscript. All authors read and approved the final manuscript.

CONFLICT OF INTEREST

Olav Lilleholt Schjørring and Thomas Lass Klitgaard were coordinating investigators of the Handling Oxygenation Targets in the Intensive Care Unit (HOT-ICU) trial,[12] and Frederik Mølgaard Nielsen is the coordinating investigator of the amended trial in COVID-19 patients (HOT-COVID).[60] Bodil Steen Rasmussen is sponsor and principal investigator of the HOT-ICU and HOT-COVID trials. Long-term outcomes (i.e., cognitive function evaluation and lung function tests) at 1-year follow-up are prespecified secondary outcomes in both trials.[23,60,61] Other authors declare no conflict of interest.

ORCID

Elena Crescioli https://orcid.org/0000-0002-8267-7634
Kirsten Uldal Krejberg https://orcid.org/0000-0001-9103-5214
Thomas Lass Klitgaard https://orcid.org/0000-0002-8781-1206
Frederik Mølgaard Nielsen https://orcid.org/0000-0002-0071-1203
Marija Barbateskovic https://orcid.org/0000-0001-8566-3660
Conni Skrubbeltrang https://orcid.org/0000-0002-7478-8422
Morten Hylander Møller https://orcid.org/0000-0002-6378-9673
Olav Lilleholt Schjørring https://orcid.org/0000-0002-7749-6003
Bodil Steen Rasmussen https://orcid.org/0000-0003-2190-145X

REFERENCES

1. Zimmerman JE, Kramer AA, Knaus WA. Changes in hospital mortality for United States intensive care unit admissions from 1988 to 2012. *Crit Care*. 2013;17:R81.
2. Desai SV, Law TJ, Needham DM. Long-term complications of critical care. *Crit Care Med*. 2011;39:371-379.
3. Schelling G, Stoll C, Vogelmeier C, et al. Pulmonary function and health-related quality of life in a sample of long-term survivors of the acute respiratory distress syndrome. *Intensive Care Med*. 2000;26: 1304-1311.
4. Herridge MS, Tansey CM, Matté A, et al. Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med*. 2011;364: 1293-1304.
5. Stevens RD, Marshall SA, Cornblath DR, et al. A framework for diagnosing and classifying intensive care unit-acquired weakness. *Crit Care Med*. 2009;37:S299-S308.
6. O'Driscoll BR, Smith R. Oxygen use in critical illness. *Respir Care*. 2019;64:1293-1307.
7. Suzuki S, Eastwood GM, Peck L, Glassford NJ, Bellomo R. Current oxygen management in mechanically ventilated patients: a prospective observational cohort study. *J Crit Care*. 2013;28:647-654.
8. 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: 1583-1589.
9. 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. *Lancet Respir Med*. 2017;5:180-190.

10. 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.
11. Mackle D, Bellomo R, Bailey M, et al. Conservative oxygen therapy during mechanical ventilation in the ICU. *N Engl J Med*. 2020;382: 989-998.
12. Schjørring OL, Klitgaard TL, Perner A, et al. Lower or higher oxygenation targets for acute hypoxemic respiratory failure. *N Engl J Med*. 2021;384:1301-1311.
13. Gelissen H, de Grooth HJ, 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*. 2021;326:940-948.
14. Barbateskovic M, Schjorring OL, Russo Krauss S, et al. Higher versus lower fraction of inspired oxygen or targets of arterial oxygenation for adults admitted to the intensive care unit. *Cochrane Database Syst Rev*. 2019;2019:CD012631.
15. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002;166:111-117.
16. Crescioli E, Uldal Krejberg K, Lass Klitgaard T, et al. The long-term effects of lower versus higher oxygenation levels in adult ICU patients - protocol for a systematic review. *Acta Anaesthesiol Scand*. 2022;66:145-151.
17. Higgins JPT, Lasserson T, Chandler J, Tovey D, Thomas J, Flemyng E, Churchill R. Methodological Expectations of Cochrane Intervention Reviews. 2021.
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
19. Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions version 6.2*. Cochrane; 2021.
20. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
21. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ*. 2003;327:557-560.
22. Barbateskovic M, Koster TM, Eck RJ, et al. A new tool to assess clinical diversity in meta-analyses (CDIM) of interventions. *J Clin Epidemiol*. 2021;135:29-41.
23. 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 Anaesthesiol Scand*. 2021;66:282-287.
24. Borland CDR, Hughes JMB. Lung diffusing capacities (D_L) for nitric oxide (NO) and carbon monoxide (CO): the evolving story. *Compr Physiol*. 2019;10:73-97.
25. Deeks J, Higgins J. Statistical Algorithms in Review Manager 5. 2007.
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials*. 1986;7:177-188.
27. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22: 719-748.
28. Demets DL. Methods for combining randomized clinical trials: strengths and limitations. *Stat Med*. 1987;6:341-350.
29. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol*. 2014;14:120.
30. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol*. 2017; 17:39.
31. Pickard AS, Neary MP, Cella D. Estimation of minimally important differences in EQ-5D utility and VAS scores in cancer. *Health Qual Life Outcomes*. 2007;5:70.
32. Della Patrona S, Zanini A, Aiello M, et al. Estimation of minimum clinically important difference in EQ-VAS score after pulmonary rehabilitation in COPD patients. *Eur Respir J*. 2014;44:P3669.

33. Zanini A, Aiello M, Adamo D, et al. Estimation of minimal clinically important difference in EQ-5D visual analog scale score after pulmonary rehabilitation in subjects with COPD. *Respir Care*. 2015;60:88-95.

34. Schünemann H, Brożek J, Guyatt G, Oxman A. GRADE Handbook for Grading Quality of Evidence and Strength of Recommendations. 2013. Accessed March 1, 2022. www.guidelinedevelopment.org/handbook

35. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011;64(4):401-406.

36. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924-926.

37. Gomersall CD, Joynt GM, Freebairn RC, Lai CK, Oh TE. Oxygen therapy for hypercapnic patients with chronic obstructive pulmonary disease and acute respiratory failure: a randomized, controlled pilot study. *Crit Care Med*. 2002;30:113-116.

38. Mazdeh M, Taher A, Torabian S, Seifirad S. Effects of normobaric hyperoxia in severe acute stroke: a randomized controlled clinical trial study. *Acta Med Iran*. 2015;53:676-680.

39. Taher A, Pilehvari Z, Poorolajal J, Aghajanloo M. Effects of normobaric hyperoxia in traumatic brain injury: a randomized controlled clinical trial. *Trauma Mon*. 2016;21:e26772.

40. Jakkula P, Reinikainen M, Hästbacka J, 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 Med*. 2018;44:2112-2121.

41. Brandt J, Spencer M, Folstein M. The Telephone Interview for Cognitive Status. *Neuropsychiatry Neuropsychol Behav Neurol*. 1988;1:111-118.

42. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20:1727-1736.

43. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics*. 2018;36:675-697.

44. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. *Md State Med J*. 1965;14:61-65.

45. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19:604-607.

46. Jennett B, Snoek J, Bond MR, Brooks N. Disability after severe head injury: observations on the use of the Glasgow outcome scale. *J Neurol Neurosurg Psychiatry*. 1981;44:285-293.

47. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet*. 1975;1:480-484.

48. Jacobs I, Nadkarni V, Bahr J, et al. Cardiac arrest and cardiopulmonary resuscitation outcome reports: update and simplification of the Utstein templates for resuscitation registries: a statement for healthcare professionals from a task force of the international liaison committee on resuscitation (American Heart Association, European resuscitation council, Australian resuscitation council, New Zealand resuscitation council, Heart and Stroke Foundation of Canada, InterAmerican Heart Foundation, resuscitation councils of southern Africa). *Circulation*. 2004;110:3385-3397.

49. Ishii K, Morimatsu H, Hyodo T, et al. 1099: relationship between inspired oxygen concentrations and atelectasis formation after extubation. *Crit Care Med*. 2018;46:533.

50. 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. *Chest*. 2019;156:A958.

51. Martin DS, McNeil M, Brew-Graves C, et al. A feasibility randomised controlled trial of targeted oxygen therapy in mechanically ventilated critically ill patients. *J Intensive Care Soc*. 2021;22:280-287.

52. 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.

53. Yang X, Shang Y, Yuan S. Low versus high pulse oxygen saturation directed oxygen therapy in critically ill patients: a randomized controlled pilot study. *J Thorac Dis*. 2019;11:4234-4240.

54. 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;33:1069-1073.

55. Lång M, Skrifvars MB, Siironen J, et al. A pilot study of hyperoxemia on neurological injury, inflammation and oxidative stress. *Acta Anaesthesiol Scand*. 2018;62:801-810.

56. Hopkins RO, Weaver LK, Pope D, Orme JF, Bigler ED, Larson LV. Neuropsychological sequelae and impaired health status in survivors of severe acute respiratory distress syndrome. *Am J Respir Crit Care Med*. 1999;160:50-56.

57. Mikkelsen ME, Christie JD, Lanken PN, et al. The adult respiratory distress syndrome cognitive outcomes study: long-term neuropsychological function in survivors of acute lung injury. *Am J Respir Crit Care Med*. 2012;185:1307-1315.

58. Angus DC, Carlet J. Surviving intensive care: a report from the 2002 Brussels roundtable. *Intensive Care Med*. 2003;29:368-377.

59. Roozenbeek B, Lingsma HF, Perel P, et al. The added value of ordinal analysis in clinical trials: an example in traumatic brain injury. *Crit Care*. 2011;15:R127.

60. Mølgaard Nielsen F, Lass Klitgaard T, Crescioli E, et al. Handling oxygenation targets in ICU patients with COVID-19-protocol and statistical analysis plan in the HOT-COVID trial. *Acta Anaesthesiol Scand*. 2021;65:1497-1504.

61. 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;63:956-965.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.