JAMA | Original Investigation | CARING FOR THE CRITICALLY ILL PATIENT

Effect of Low-Normal vs High-Normal Oxygenation Targets on Organ Dysfunction in Critically Ill Patients
A Randomized Clinical Trial

Harry Gelissen, MD, MBA; Harm-Jan de Grooth, MD, PhD; Yvo Smulders, MD, PhD; Evert-Jan Wils, MD, PhD; Wouter de Ruijter, MD, PhD; Roel Vink, MD, PhD; Bob Smit, PhD; Jantine Röttgering, MD; Leila Atmowihardjo, MD; Armand Girbes, MD, PhD; Paul Elbers, MD, PhD; Pieter-Roel Tuinman, MD, PhD; Heleen Oudemans-van Straaten, MD, PhD; Angelique de Man, MD, PhD

➕ **Visual Abstract**

➕ **Editorial**

➕ **Supplemental content**

IMPORTANCE Hyperoxemia may increase organ dysfunction in critically ill patients, but optimal oxygenation targets are unknown.

OBJECTIVE To determine whether a low-normal PaO_2 target compared with a high-normal target reduces organ dysfunction in critically ill patients with systemic inflammatory response syndrome (SIRS).

DESIGN, SETTING, AND PARTICIPANTS Multicenter randomized clinical trial in 4 intensive care units in the Netherlands. Enrollment was from February 2015 to October 2018, with end of follow-up to January 2019, and included adult patients admitted with 2 or more SIRS criteria and expected stay of longer than 48 hours. A total of 9925 patients were screened for eligibility, of whom 574 fulfilled the enrollment criteria and were randomized.

INTERVENTIONS Target PaO_2 ranges were 8 to 12 kPa (low-normal, n = 205) and 14 to 18 kPa (high-normal, n = 195). An inspired oxygen fraction greater than 0.60 was applied only when clinically indicated.

MAIN OUTCOMES AND MEASURES Primary end point was $SOFA_{RANK}$, a ranked outcome of nonrespiratory organ failure quantified by the nonrespiratory components of the Sequential Organ Failure Assessment (SOFA) score, summed over the first 14 study days. Participants were ranked from fastest organ failure improvement (lowest scores) to worsening organ failure or death (highest scores). Secondary end points were duration of mechanical ventilation, in-hospital mortality, and hypoxemic measurements.

RESULTS Among the 574 patients who were randomized, 400 (70%) were enrolled within 24 hours (median age, 68 years; 140 women [35%]), all of whom completed the trial. The median PaO_2 difference between the groups was −1.93 kPa (95% CI, −2.12 to −1.74; P < .001). The median $SOFA_{RANK}$ score was −35 points in the low-normal PaO_2 group vs −40 in the high-normal PaO_2 group (median difference, 10 [95% CI, 0 to 21]; P = .06). There was no significant difference in median duration of mechanical ventilation (3.4 vs 3.1 days; median difference, −0.15 [95% CI, −0.88 to 0.47]; P = .59) and in-hospital mortality (32% vs 31%; odds ratio, 1.04 [95% CI, 0.67 to 1.63]; P = .91). Mild hypoxemic measurements occurred more often in the low-normal group (1.9% vs 1.2%; median difference, 0.73 [95% CI, 0.30 to 1.20]; P < .001). Acute kidney failure developed in 20 patients (10%) in the low-normal PaO_2 group and 21 patients (11%) in the high-normal PaO_2 group, and acute myocardial infarction in 6 patients (2.9%) in the low-normal PaO_2 group and 7 patients (3.6%) in the high-normal PaO_2 group.

CONCLUSIONS AND RELEVANCE Among critically ill patients with 2 or more SIRS criteria, treatment with a low-normal PaO_2 target compared with a high-normal PaO_2 target did not result in a statistically significant reduction in organ dysfunction. However, the study may have had limited power to detect a smaller treatment effect than was hypothesized.

TRIAL REGISTRATION ClinicalTrials.gov Identifier: NCT02321072

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Angelique de Man, MD, PhD, Department of Intensive Care Medicine, Research VUmc Intensive Care, Amsterdam Cardiovascular Science, Amsterdam Infection and Immunity Institute, Amsterdam Medical Data Science, Amsterdam UMC, location VUmc, De Boelelaan 1117, 1007MB Amsterdam, the Netherlands (ame.deman@amsterdamumc.nl).

Section Editor: Christopher Seymour, MD, Associate Editor, *JAMA* (christopher.seymour@jamanetwork.org).

O xygen is often liberally administered in the intensive care unit (ICU) to treat or prevent hypoxemia. For many years, oxygen was considered to have consistently favorable effects. However, an observational study using data from 1999 to 2006 showed a U-shaped relationship between PaO_2 and mortality,[1] opening a debate about optimal oxygenation targets in critically ill patients.

The potential negative effects of hyperoxemia include pulmonary toxicity,[2] augmented ischemia/reperfusion injury, and systemic vasoconstriction with decreased organ perfusion.[3] These effects may impair rather than improve tissue oxygen delivery.[2,4] Conversely, hyperoxemia may also have benefit. Systemic vasoconstriction may curtail vasodilation in patients with systemic inflammatory response syndrome[5] (SIRS) and favorably redistribute blood flow to organs.[3] Hyperoxemia may also have antimicrobial effects.[6]

Six randomized clinical trials (RCTs) in ICU patients[7-12] and 1 meta-analysis[13] comparing low vs high oxygenation targets showed inconsistent results. Two of these trials[7,8] and the meta-analysis[13] reported reduced mortality for lower oxygenation targets, while 4 trials[9-12] showed no difference in outcome. One trial[11] reported an increased incidence of mesenteric ischemia in the low oxygenation target group.

There are limitations to these trials. First, none specifically enrolled patients with systemic inflammation, who might benefit most from the postulated vasoconstrictive effects of hyperoxemia. Second, no trial defined an upper limit for inspired oxygen fraction (FIO_2) to attain higher oxygen saturation as measured by pulse oximetry (SpO_2) targets, thereby potentially exposing patients to possible pulmonary oxygen toxicity. Third, 4 RCTs[7-10] used SpO_2-based targets for oxygenation instead of PaO_2 targets, a more accurate strategy.[14]

The aim of this trial was to investigate whether a low-normal PaO_2 target compared with a high-normal PaO_2 target (avoiding toxic FIO_2 values) in critically ill patients with systemic inflammation leads to improved organ function.

Methods

Trial Design

This multicenter RCT was conducted in 1 academic hospital (Amsterdam UMC, location VUmc) and 3 nonacademic hospitals (Tergooiziekenhuizen, Noordwest Ziekenhuisgroep, and Franciscus Gasthuis & Vlietland) in the Netherlands. The medical ethical committee of Amsterdam UMC, location VUmc, approved the protocol including the statistical analysis plan of the study (METc No. 2014.459) (available in Supplement 1). The Clinical Research Office of Amsterdam UMC, location VUmc, monitored the study. Written deferred consent was obtained within 1 day from admission from patients or their legal representatives.

Patients

Patients aged 18 years or older admitted to the ICU with an expected stay of 48 hours or longer and systemic inflammation (defined as ≥2 positive SIRS criteria[5]) were eligible for the trial. Exclusion criteria were (1) admission after elective

Key Points

Questions Does a low-normal oxygenation compared with a high-normal oxygenation target range reduce organ dysfunction in critically ill patients?

Findings This randomized clinical trial included 400 patients in the intensive care unit with at least 2 positive systemic inflammatory response syndrome criteria. Randomization to a target PaO_2 range of 8 to 12 kPa vs 14 to 18 kPa resulted in a median $SOFA_{RANK}$ score of −35 vs −40 (lower score represents less organ failure severity), a difference that was not statistically significant.

Meanings Among critically ill patients, targeting oxygenation to a low-normal range compared with a high-normal range did not result in a statistically significant reduction in organ dysfunction.

surgery, (2) known pulmonary arterial hypertension World Health Organization class III or IV, (3) severe acute respiratory distress syndrome (ARDS; according to the Berlin criteria), (4) pregnancy, (5) severe chronic obstructive pulmonary disease (Global Initiative for Chronic Obstructive Lung Disease class III/IV), (6) a do-not-intubate order, (7) carbon monoxide or cyanide intoxication or methemoglobinemia, (8) sickle cell disease, and (9) known cardiac right-to-left shunting.

Randomization

Patients were screened for eligibility immediately on ICU admission. Eligible patients were randomized to either a low-normal or a high-normal PaO_2 target range in a 1:1 ratio using web-based randomization with randomly permuted blocks of size 4 to 8, stratified by age (<50, 50-70, and >70 years), sex, and reason of admittance (medical, surgical, or trauma).

Study Intervention

Oxygen was administered to target either a low-normal PaO_2 range of 8 to 12 kPa or a high-normal PaO_2 range from 14 to 18 kPa. The target PaO_2 was achieved by adjustment of the FIO_2 (administered through a mechanical ventilator, nasal cannula, oxygen mask, or high-flow oxygen therapy) and/or positive end-expiratory pressure (PEEP) if patients were mechanically ventilated. Levels to achieve the target PaO_2 were maximized to an FIO_2 of 0.60 and 10 cm H_2O of PEEP (eFigure 1 in Supplement 2) unless clinically otherwise indicated. Prone positioning during mechanical ventilation, change from oxygen mask to noninvasive mechanical ventilation, and intubation were only applied on clinical indication, not to achieve the PaO_2 targets of the trial. Temporary measures to increase oxygenation during planned procedures involving upper airways, such as tracheostomy or bronchoscopy, were used according to standard practices at participating centers. These deviations from study targets were limited to the shortest duration possible. Arterial blood samples were drawn according to routine clinical practice in all participating hospitals. When a patient was not in target range, oxygen administration was adjusted and an arterial blood sample was taken after 15 minutes to check whether the patient was in target range.

Figure 1. Patient Selection, Randomization, and Flow Through the Trial
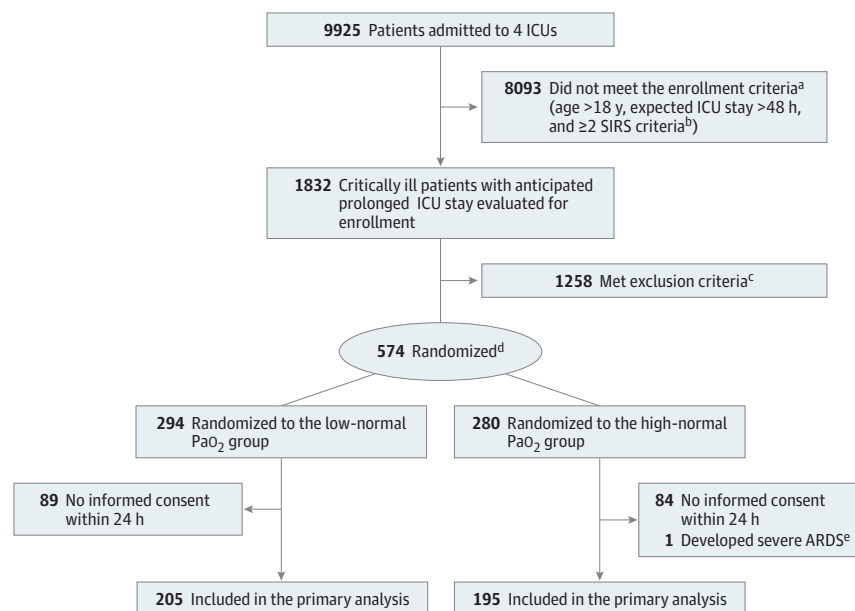
ICU indicates intensive care unit.

a Screening details were not reported by all 4 participating centers. For the 2 centers that did report screening data, results are available in the eFigure 2 in Supplement 2.

b SIRS criteria: (1) temperature >38 °C or <36 °C; (2) heart rate >90 beats/min; (3) respiratory rate >20 breaths/min or $PaCO_2$ <32 mm Hg; or (4) white blood cell count >12 ×10^9/L, <4 ×10^9/L, or >10% immature (band) forms.

c Exclusion criteria: admitted after elective surgery, known pulmonary arterial hypertension, severe acute respiratory distress syndrome (ARDS), pregnancy,

severe chronic obstructive pulmonary disease, do-not-intubate order, carbon monoxide or cyanide intoxication, methemoglobinemia, sickle cell disease, or known cardiac right-to-left shunting. Reasons for exclusion were not reported by all participating centers. Data for the reporting centers are available in eFigure 2 in Supplement 2.

d Randomization was stratified in a 1:1 ratio by age (<50, 50-70, and >70 years), sex, and reason of admittance (medical, surgical, or trauma).

e This patient developed severe ARDS immediately after randomization.

The intervention commenced within 12 hours after ICU admission and was maintained until the earliest occurrence of either day 14, ICU discharge, or death. Time-weighted values of PaO_2, arterial oxygen saturation (SaO_2), SpO_2, and FIO_2 were calculated. Data collection is described in eAppendix 1 in Supplement 2.

Primary End Point

The primary end point was a ranking based on the nonrespiratory cumulative daily delta Sequential Organ Failure Assessment (SOFA) score from day 1 to day 14 ($SOFA_{RANK}$).

The daily SOFA score was calculated as the total of maximum scores for each organ system, excluding the respiratory system because of likely PaO_2/FIO_2 distortion.[15,16] For each patient, the daily total SOFA score minus the baseline SOFA score was summed over the first 14 study days. Discharge was counted (from the day of discharge forward) as a score of 0 minus baseline score, and death was counted (from the day of death forward) as a maximum score of 20 minus baseline score. The resulting cumulative daily delta score was used to rank participants from fast organ failure improvement (lowest scores) to worsening organ failure and death (highest scores). Given a baseline score, early organ function improvement resulted in a lower (better) score than late improvement, which in turn resulted in a lower score than deteriorating organ function. An outcome of death resulted in the

highest scores, but death after a high baseline score resulted in a better score than death after a low baseline score. Death was given the maximum score to prevent survivorship bias.[17]

The absolute values and relative changes in the primary end point are not directly interpretable to clinicians. Instead, the primary end point was designed to correlate with mortality and to respond to treatments that affect mortality.[18] In all, $SOFA_{RANK}$ is a composite reflection of organ failure burden, mortality, and length of stay over 14 days, balanced against the baseline organ dysfunction at trial enrollment. Example calculations of the primary end point are shown in the eAppendix 2 in Supplement 2.

Secondary End Points

Because the primary end point was novel, the predefined secondary end points included 3 other SOFA score–derived outcomes as a sensitivity analysis of the design choices of the primary end point. These secondary SOFA score end points were the maximum nonrespiratory SOFA score, maximum nonrespiratory SOFA score minus baseline score, and the nonrespiratory SOFA score rate of decline.

Other predefined secondary end points were the number of hypoxemic measurements (defined as mild: PaO_2, 5-7.3 kPa and severe: PaO_2 <5 kPa), time spent in the assigned PaO_2 range, duration of mechanical ventilation and ventilator-free days

Table 1. Baseline Characteristics of the Patients

Groups	No. (%)	
	Low-normal PaO$_2$ target (n = 205)	High-normal PaO$_2$ target (n = 195)
Age, median (IQR), y	68 (56-76)	68 (61-75)
Sex		
Male	134 (65)	126 (65)
Female	71 (35)	69 (35)
Reason for admission		
Medical	143 (70)	139 (71)
Surgical	50 (24)	41 (21)
Trauma	12 (6)	15 (8)
Chronic diagnoses[a]		
Diabetes	41 (20)	43 (22)
Immune compromised	25 (12)	22 (11)
Chronic kidney failure not requiring kidney replacement therapy	17 (8.3)	18 (9.3)
Cancer	17 (8.3)	15 (7.7)
COPD (drug dependent)	15 (7.3)	20 (10)
Hematological malignancy	10 (4.9)	10 (5.1)
Recent (<6 mo) myocardial infarction	10 (4.9)	12 (6.2)
Kidney replacement therapy	6 (2.9)	7 (3.6)
NYHA IV heart failure[b]	4 (1.9)	3 (1.5)
Liver cirrhosis	3 (1.5)	4 (2.1)
Acute diagnoses		
Systemic infection[c]	71 (35)	73 (38)
Pneumonia	69 (34)	59 (30)
Cardiac arrest	37 (18)	42 (22)
Acute abdominal infection/infarction	34 (17)	30 (15)
Major bleeding	24 (12)	17 (8.8)
Soft tissue infection	12 (5.8)	4 (2.1)
Stroke	5 (2.4)	6 (3.1)
Oxygenation[d]		
PaO$_2$, median (IQR), kPa	11.6 (9.5-16)	12.3 (10.4-14.8)
FiO$_2$, median (IQR)	45 (39-60)	46 (40-60)
Invasive or noninvasive mechanical ventilation	153 (75)	142 (73)
Positive end-expiratory pressure, cm H$_2$O[e]		
No.	153	142
Median (IQR)	8 (5-10)	8 (5-10)
Additional measures and treatments, median (IQR)		
Admission SOFA score (excluding the respiratory component)[f]	5 (3-8)	6 (4-8)
Vasopressor use at enrollment	152 (71.1)	152 (77.9)
Time from ICU admission to randomization, h	4.0 (1.5-7.1)	4.0 (2.1-7.0)

Abbreviations: COPD, chronic obstructive pulmonary disease; FiO$_2$, fraction of inspired oxygen; ICU, intensive care unit; IQR, interquartile range; NYHA, New York Heart Association; SOFA, Sequential Organ Failure Assessment.

SI conversion factor: To convert PaO$_2$ to mm Hg, divide by 0.133.

[a] More than 1 chronic diagnosis could apply to the same patient.

[b] NYHA class IV: cardiac disease with symptoms (such as fatigue, palpitations, dyspnea, and/or angina) occurring at rest.

[c] Clinical signs of generalized infection combined with positive blood cultures.

[d] Initial oxygen measures were taken at the time of enrollment.

[e] Only used in patients who were ventilated.

[f] The SOFA score is a cumulative score of the respiratory system (PaO$_2$/FiO$_2$), nervous system (Glasgow Coma Scale), cardiovascular system (mean arterial pressure or vasopressor use), liver (bilirubin), coagulation (platelets), and kidney function (creatinine or urine output) used to assess the severity of organ failure in the ICU. Each system scores 0 to 4 for a total of 0 to 24. If the respiratory system is excluded, the maximum score is 20.

until day 14, vasopressor use until day 14, fluid balances until day 14, F2-isoprostanes, ICU and hospital length of stay, ICU mortality, and in-hospital mortality. Post hoc–defined secondary end points were 90-day mortality, and the number of hyperoxic measurements, defined using 3 cutoffs: 13.3 kPa, 16.7 kPa (within the range of high-normal PaO$_2$), and 18 kPa (above the highest value of high-normal PaO$_2$).

Adverse Events

The following serious adverse events were recorded: new myocardial infarction during ICU admission, stroke occurring during the ICU admission, kidney failure with need for kidney replacement therapy more than 24 hours after ICU admission, need for prone positioning, and new severe liver failure. The definitions of the serious adverse events can be found in eAppendix 3 in Supplement 2.

Statistical Analysis

The trial was designed to detect a difference of 0.33 SDs on the primary end point with 90% power and a 2-sided α of .05 with 2 interim analyses. In a validation study of SOFA score–based end points, a between-group difference in SOFA trajectory of 0.33 SDs was associated with a 28-day mortality odds ratio of 0.79 (e^[−0.33 × 0.70]) (95% CI, 0.69-0.92).[18] Thus, the total

Figure 2. PaO_2 by Treatment Group for Patients Alive in the Intensive Care Unit Days 1 to 15
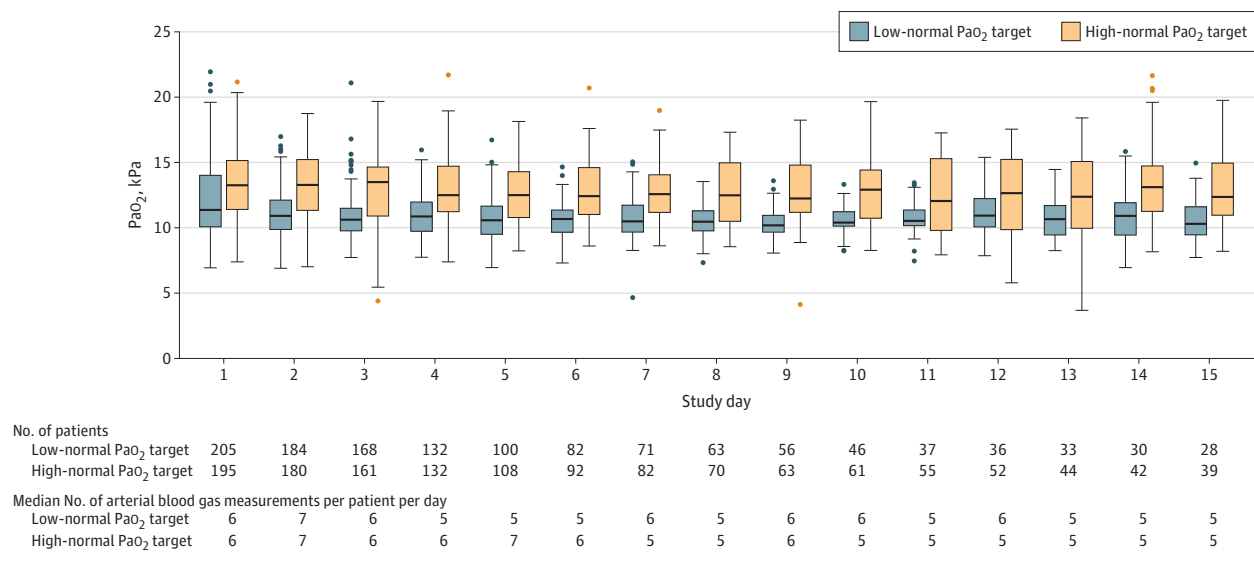
Boxes represent medians and interquartile ranges, whiskers extend to the lowest and highest observations within 1.5 × interquartile range (IQR), and the circles represent outlier observations. Plotted values were based on a median of 5 (IQR, 3 to 7) arterial blood gas measurements per patient per day, which were time-weighted and averaged per patient-day before aggregation by study group. The median PaO_2 difference between the study groups was −1.93 kPa (95% CI, −2.12 to −1.74; $P < .001$), with significant differences on each study day except day 12 ($P = .07$). Boxes are offset for readability. To convert PaO_2 to mm Hg, divide by 0.133.

sample size proposed for this trial was 385 patients. At a minimum asymptotic relative efficiency of the Wilcoxon rank-sum test compared with a t test, the power to detect an effect with this sample size was 85%.[19]

Interim analyses were conducted after inclusion of 150 and 275 patients with predefined stopping rules for outcome differences in either direction. Alpha spending for the interim analyses was approximated with a Lan-DeMets and O'Brien-Fleming spending function.[20] The predefined level of significance for the final analysis of the primary end point was a P value of .0452 (accounting for alpha spending on the 2 interim analyses). Because of the potential for type I error due to multiple comparisons, findings for analyses of secondary end points should be interpreted as exploratory.

Statistical analyses were performed using the SPSS statistical software package (SPSS Inc) and the R language and environment for statistical computing (R Foundation for Statistical Computing) with the *tidyverse* (https://tidyverse. tidyverse.org/) and *survival* suite of packages.[21]

Continuous normally distributed data were reported as means with SDs. Other continuous data were reported as medians and interquartile ranges (IQRs). The treatment groups were compared according to their randomization group. All patients for whom deferred consent was obtained within 24 hours were included in the analyses. There were no missing outcome data. The primary end point was designed as a tool for ranking outcomes, and the between-group comparisons were conducted using the Wilcoxon rank-sum test. As a post hoc analysis, the primary outcome was adjusted for the effect by the stratification variables: site, reason of admittance, age category, and sex.[22,23] The stratification variables were entered

as random effects in a mixed-effects model with treatment allocation as fixed effect. Secondary outcomes were compared using Fisher exact test (reported as odds ratios) and Wilcoxon rank-sum tests (reported as median differences). Mortality over time was assessed with Kaplan-Meier curves and compared using a Cox proportional hazards model. The proportional hazards assumption was evaluated by testing the independence between model residuals and time and was found to be satisfied ($P = .76$ for 90-day mortality).

Results

Patients

From February 2015 until January 2019, 9925 patients were screened for eligibility. Among these, 574 met enrollment criteria and were randomized. Detailed screening data were available at 2 centers (eFigure 2 in Supplement 2). Informed consent was obtained for 400 participants (205 in the low-normal PaO_2 group and 195 in the high-normal PaO_2 group; **Figure 1**). No patients were lost to follow-up and there were no missing baseline or end point data. Baseline characteristics were comparable in both groups (**Table 1**). The median time from ICU admission to inclusion was 4.0 hours (IQR, 1.5-7.1) in the low-normal PaO_2 group and 4.0 hours (IQR, 2.1-7.0) in the high-normal PaO_2 group.

Oxygenation

The median PaO_2 in the low-normal group (10.8 kPa [IQR, 9.8-12.0 kPa]) was significantly lower compared with the high-normal group (12.8 kPa [IQR, 10.9-14.9 kPa]). The time-weighted median SaO_2, SpO_2, and FIO_2 were significantly

Figure 3. Cumulative SOFA$_{RANK}$ Outcomes by Treatment Group
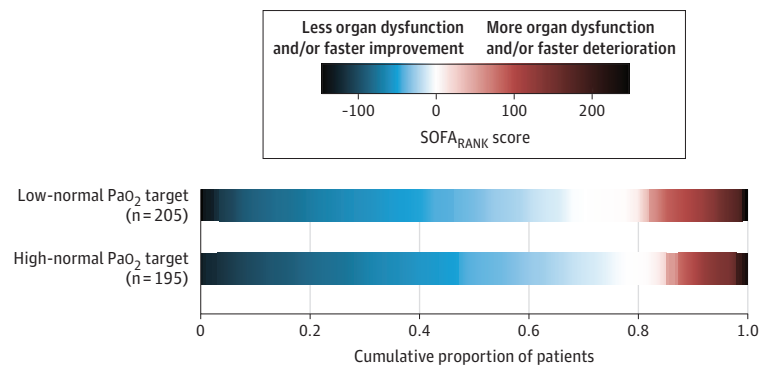
The primary end point was a ranked outcome of nonrespiratory organ failure quantified by the nonrespiratory components of the Sequential Organ Failure Assessment (SOFA) score. For each patient, the daily SOFA score minus the baseline SOFA score was summed over the first 14 study days. The resulting score was used to rank participants from fastest organ failure improvement (lowest scores) to worsening organ failure or death (highest scores).

The figure shows that patients randomized to a high-normal PaO$_2$ target had nonsignificantly lower (better) scores. The median SOFA$_{RANK}$ score was −35 points (interquartile range [IQR], −63 to 0) in the low-normal group vs −40 points (IQR, −76 to −4.5) in the high-normal group (median difference, 10 points [95% CI, 0 to 21]; *P* = .06).

lower in the low-normal group (*P* < .001 for all). Mild hypoxemic measurements were significantly more common in the low-normal PaO$_2$ group (1.9% vs 1.2%; median difference, 0.73 [95% CI, 0.30-1.20]; *P* < .001), but there was no significant difference in severe hypoxemia (eTable 1 in Supplement 2). Hyperoxemic measurements were significantly more common in the high-normal PaO$_2$ group.

The median PaO$_2$ values in both groups on days 1 to 15 are provided in **Figure 2** and were significantly lower in the low-normal group on all days, except day 12. The median number of arterial blood samples per patient per day slightly decreased from 6 on day 2 to 4 on day 15 (eTable 2 in Supplement 2).

Primary End Point

The median SOFA$_{RANK}$ score was −35 points (IQR, −63 to 0) in the low-normal group vs −40 points (IQR, −76 to −4.5) in the high-normal group (median difference, 10 [95% CI, 0 to 21]; *P* = .06; **Figure 3**; eFigure 3 in Supplement 2). Adjustment for the stratification variables did not change the treatment effect on the primary end point (eAppendix 4 in Supplement 2).

Secondary End Points

SOFA score–related secondary end points were not significantly different between the low-normal and the high-normal PaO$_2$ groups (**Table 2**). Both adjusted and unadjusted SOFA$_{RANK}$ measures demonstrated no significant difference between the treatment groups (eFigure 4 in Supplement 2). To illustrate the potential effect of survivorship bias, results are split in crude data showing the daily SOFA scores and the number of participants still in the ICU at days 1 to 15 in eFigure 4A in Supplement 2, and data adjusted for death and discharge for all participants at days 1 to 15 in eFigure 4B in Supplement 2. There were no significant differences between the low-normal and the high-normal PaO$_2$ groups for the other non-SOFA-related secondary end points (Table 2). There were no significant differences in the duration of mechanical ventilation (eFigure 5 in Supplement 2), length of ICU stay (eFig-

ure 6 in Supplement 2), or mortality from inclusion until day 90 (eFigure 7 in Supplement 2).

Adverse Events

Acute kidney failure developed in 20 patients (10%) in the low-normal PaO$_2$ group and 21 patients (11%) in the high-normal PaO$_2$ group and acute myocardial infarction in 6 patients (2.9%) in the low-normal PaO$_2$ group and 7 patients in the high-normal PaO$_2$ group (3.6%) (Table 2).

Discussion

In this multicenter RCT including 400 critically ill patients with 2 or more SIRS criteria, treatment with a low-normal PaO$_2$ target (8-12 kPa) compared with a high-normal target (14-18 kPa) did not significantly reduce organ dysfunction at 14 days. There were no significant differences in 90-day mortality, duration of mechanical ventilation, or ICU length of stay.

The high-normal PaO$_2$ targets were based on current clinical practice and avoidance of potential detrimental effects of severe hyperoxemia. The FIO$_2$ to attain the high-normal target was restricted to a maximum of 0.60. By comparison, earlier trials did not report an upper limit for FIO$_2$ to achieve the high oxygenation target. Similar to 2 other trials,[11,12] we evaluated oxygen therapy predominantly based on PaO$_2$ targets. A PO$_2$ difference is the driving force for oxygen diffusion from the arterial blood to the tissues.[14] Small differences in SpO$_2$ can coincide with large PaO$_2$ differences because the relation between PaO$_2$ and SpO$_2$ is not linear but S-shaped. Therefore, a PaO$_2$ strategy allows more precise titration of oxygenation compared with SpO$_2$ targets, although with a theoretical need for increased blood sampling.[24-26]

Previous trials showed either more favorable outcomes for the low oxygenation groups[7,8,13] or showed no significant differences between low and high oxygenation targets.[9-12] Several differences between the trials are noteworthy. The 2 trials

Table 2. Primary (SOFA-Based Ranking) and Secondary End Points and Adverse Events

Group	Low-normal Pao$_2$ target (n = 205)	High-normal Pao$_2$ target (n = 195)	Difference (95% CI)[a]	Odds ratio (95% CI)	P value
Primary end point					
SOFA$_{RANK}$ (score over 14 d), median (IQR)[b]	−35 (−63 to 0)	−40 (−76 to −4.5)	10 (0 to 21)		.06
Secondary end points[c]					
Maximum nonrespiratory SOFA score, median (IQR)[d]	6 (3 to 12)	6 (4 to 9.5)	0 (−1 to 0)		.41
Maximum nonrespiratory SOFA score minus baseline score, median (IQR)[d]	0 (−1 to 2)	0 (−1 to 1.5)	0 (−1 to 1)		.46
SOFA score rate of decline, mean (SD), points/d[d]	−0.28 (0.22)	−0.28 (0.29)	0.01 (−0.03 to 0.05)		.78
Intubation during ICU admission, No. (%)[e]	147 (72)	143 (73)	−1 (−11 to 8)	0.92 (0.58 to 1.46)	.74
Duration of mechanical ventilation, median (IQR), d	3.4 (1.2 to 6.8)	3.1 (1.4 to 9.7)	−0.15 (−0.88 to 0.47)[f]		.59
Ventilator-free days to day 14, median (IQR)	9.7 (0 to 13.4)	10.2 (0 to 13.5)	0 (−1.2 to 1.3)		.85
Length of stay in ICU, median (IQR), d	3.9 (2.0 to 8.3)	4.6 (2.0 to 11.1)	−0.34 (−1.14 to 0.37)		.34
Total norepinephrine dose during admission, median (IQR), mg	27.6 (0.7 to 97.2)	27.6 (7.4 to 76.8)	0 (−7.2 to 4.8)		.74
Highest norepinephrine dose during admission, median (IQR), mg/h	0.55 (0.10 to 1.60)	0.60 (0.20 to 1.40)	0 (−0.12 to 0.10)		.77
Mortality, No. (%)					
ICU	50 (24)	49 (25)	−1 (−10 to 8)	0.96 (0.59 to 1.55)	.91
Hospital	66 (32)	61 (31)	1 (−9 to 10)	1.04 (0.67 to 1.63)	.91
At 90 d[e]	72 (35)	67 (34)	1 (−9 to 11)	1.03 (0.67 to 1.59)	.91
Causes of death, No./total (%)[g]					
Cardiac	29/72 (40)	26/67 (39)			
Systemic infection and/or multiple organ failure	28/72 (39)	24/67 (36)			
Irreversible brain injury secondary to OHCA	22/72 (31)	12/67 (18)			
Pulmonary	18/72 (25)	14/67 (21)			
Hepatic	9/72 (13)	2/67 (3)			
Brain trauma	8/72 (11)	10/67 (15)			
Severe bleeding	7/72 (10)	13/67 (19)			
Stroke	5/72 (7)	1/67 (1)			
Oncology	3/72 (4)	2/67 (3)			
Missing	0/72	1/67 (1)			
Serious adverse events, No. (%)					
Kidney replacement therapy for acute kidney failure	20 (10)	21 (11)			
Severe respiratory failure necessitating prone ventilation	7 (3.4)	8 (4.1)			
New myocardial infarction	6 (2.9)	7 (3.6)			
New liver failure	7 (3.4)	2 (1.0)			
New stroke	1 (0.5)	2 (1.0)			
No. of patients with ≥1 serious adverse events (including deaths)	89 (43)	79 (41)			

Abbreviations: ICU, intensive care unit; IQR, interquartile range; OHCA, out-of-hospital cardiac arrest; SOFA, Sequential Organ Failure Assessment.

[a] For end points reported as medians, the difference is the median difference between the high-normal vs the low-normal Pao$_2$ target group, which is not necessarily equal to the difference in reported medians. For end points reported as means or percentages, the difference is the absolute difference.

[b] The trial primary end point was a ranked outcome of nonrespiratory organ failure quantified by the nonrespiratory components of the SOFA score. For each patient, the daily total SOFA score minus the baseline SOFA score was summed over the first 14 study days. Discharge was counted (from the day of discharge forward) as a score of 0 minus baseline score and death was counted (from the day of death forward) as a maximum score of 20 minus baseline score. The resulting cumulative daily delta score was used to rank participants from fastest organ failure improvement (lowest scores) to worsening organ failure or death (highest scores).

[c] All secondary outcomes were obtained over the 14 days after randomization, except adverse events and length of stay (obtained over the entire ICU stay) and mortality (obtained over the indicated timeframes).

[d] All SOFA-based scores are calculated excluding the respiratory component. SOFA score is a cumulative score of the respiratory system (Pao$_2$/Fio$_2$), nervous system (Glasgow Coma Scale), cardiovascular system (mean arterial pressure or vasopressor use), liver (bilirubin), coagulation (platelets), and kidney function (creatinine or urine output) used to assess the severity of organ failure in the ICU. Each system scores 0 to 4 for a total of 0 to 24. If the respiratory system is excluded, the maximum score is 20.

[e] End points added post hoc.

[f] The sign of the median difference is opposite to the apparent difference in medians, which can occur in nonparametric comparisons.

[g] Registered by the local investigator in the Case Report Form.

that found more favorable outcomes for the low oxygenation strategy were both prematurely terminated. The Oxygen-ICU trial[7] was stopped after inclusion of 434 patients due to logistic factors at a moment when the low oxygenation group had significantly lower mortality. The HYPERS2S trial, in which mechanically ventilated patients with sepsis were randomized to 100% oxygen vs an SpO_2 of 88% to 95%,[8] was stopped prematurely because of excess mortality in the hyperoxemia group that did not reach statistical significance.

The 2 largest RCTs before this study did not find a significant difference in mortality between low and high oxygenation targets.[10,12] The ICU-ROX trial randomized 965 mechanically ventilated patients to different SpO_2-based oxygenation targets and found no significant differences in ventilator-free days or survival.[10] The HOT-ICU trial randomized 2928 patients with relatively severe acute hypoxemic respiratory failure to target a PaO_2 of 8 kPa vs 12 kPa. In this population with a median baseline FIO_2 of 0.70 and baseline PaO_2:FIO_2 ratio of approximately 120 mm Hg, there was no significant mortality difference between the oxygenation groups.[12]

The recent $LOCO_2$ trial[11] in patients with ARDS, with a target PaO_2 of 7.3 to 9.3 kPa vs 12 to 14 kPa, was stopped early because of 5 cases of mesenteric ischemia in the low oxygenation group, whereas none occurred in the high oxygenation group. At that point, 90-day mortality was significantly higher in the low oxygenation group. However, the much larger HOT-ICU trial did not find an important difference in serious adverse events, including mesenteric ischemia (low group: 2.2% vs high group: 2.0%).

A meta-analysis comparing low vs high oxygen therapy in acutely ill patients (mostly admitted with stroke or myocardial infarction) showed a significantly lower mortality with low oxygen therapy.[13] Most studies included in this meta-analysis[13] used fixed FIO_2 without specific oxygenation targets and the duration of the interventions was relatively short. This meta-analysis suggested that liberal oxygen therapy increases mortality in acutely ill patients, but the analysis did not identify specific oxygenation targets.

The present trial investigating critically ill patients with signs of systemic inflammation showed improved resolution of nonrespiratory organ failure in the high-normal group compared with the low-normal group, but this difference was not statistically significant. There was a small but significant increase in the incidence of hypoxemia in the low-normal PaO_2 target group, but a much larger increase in the incidence of hyperoxemia in the high-normal PaO_2 target group. In the $LOCO_2$ trial,[11] 59% of the patients in the low oxygenation group had an episode with a PaO_2 below 7.3 kPa. In the HOT-ICU trial, the number of hypoxic episodes was not reported. It cannot be ruled out that high-normal oxygenation targets might be related to improved organ function because of improved oxygen delivery at the tissue level. Alternatively, hyperoxemia-induced systemic vasoconstriction[27] might have counterbalanced SIRS-induced vasoplegia and thereby caused a favorable redistribution of circulation to vital organs, as previously shown in porcine models.[28,29] Limiting FIO_2 to 0.60 possibly minimized negative effects due to direct oxidative pulmonary toxicity.[3] Altogether, the balance of effects from hypoxemia vs hyperoxemia remains unclear. The body of trials

investigating mild hyperoxemic vs normoxemic targets, including the current study, has not clearly demonstrated an effect on outcomes. The effect of hyperoxemia is likely smaller than has been hypothesized (detecting a possibly very small effect is the goal of the currently ongoing Mega-ROX trial [UMIN Clinical Trials Registry UMIN000042551]) or an effect is present only at more extreme values of hyperoxemia.

This study has many strengths. It enrolled patients at 1 academic and 3 nonacademic ICUs. An independent quality officer frequently monitored all participating centers. The primary end point ($SOFA_{RANK}$) was novel. Although difficult for clinical interpretation, it was specifically designed to avoid the pitfalls of surrogate end points and competing risks. The "surrogate paradox" occurs when a disease-oriented outcome (such as SOFA score) fails to capture the treatment effects on a patient-oriented outcome (such as mortality) even though the 2 outcomes are strongly correlated.[30,31] SOFA trajectory, the driving component of $SOFA_{RANK}$, has been shown to capture treatment effects on mortality, avoiding this paradox. Simpler SOFA end points, such as the score on a fixed day after randomization, fail to satisfy this criterion.[18] Competing risks of mortality and discharge can render a seemingly straightforward SOFA end point invalid and make study results difficult to interpret. For example, the increased mortality can paradoxically improve the mean SOFA score in simpler iterations of the SOFA end point.[17] The most important drawback from the present approach is the fact that the absolute value of $SOFA_{RANK}$ has no directly interpretable clinical meaning. In this context, the end point was designed solely to test the hypothesis that the intervention would affect the trajectory of organ failure while explicitly taking into account the competing risks of death and discharge.

Limitations

This study had several limitations. First, short episodes of higher FIO_2 were allowed in both groups if considered necessary by the treating physician, which may have attenuated the treatment contrast between the groups. Second, PaO_2 levels before inclusion may also have reduced an effect of the oxygenation difference between both groups. Third, FIO_2 was only increased above 0.60 on clinical indication, not to achieve the target PaO_2. Although the difference between the achieved median PaO_2 values in the low-normal and the high-normal groups was statistically significant, the median PaO_2 value in the high-normal group was below the target range. This limits the inferences about the safety of the high-normal PaO_2 targets. Fourth, the study was designed to detect a 0.33-SD difference in the primary outcome, but the true effect on organ dysfunction may be smaller.

Conclusions

Among critically ill patients with at least 2 or more SIRS criteria, treatment with a low-normal PaO_2 target compared with a high-normal PaO_2 target did not result in a statistically significant reduction in organ dysfunction. However, the study may have had limited power to detect a smaller treatment effect than was hypothesized.

Author Affiliations: Department of Intensive Care Medicine, Research VUmc Intensive Care, Amsterdam Cardiovascular Science, Amsterdam Infection and Immunity Institute, Amsterdam Medical Data Science, Amsterdam UMC, location VUmc, Amsterdam, the Netherlands (Gelissen, de Grooth, Smit, Röttgering, Atmowihardjo, Girbes, Elbers, Tuinman, Oudemans-van Straaten, de Man); Department of Anesthesiology, Amsterdam UMC, location VUmc, Amsterdam, the Netherlands (de Grooth); Department of Internal Medicine, Amsterdam UMC, location VUmc, Amsterdam, the Netherlands (Smulders); Department of Intensive Care, Franciscus Gasthuis & Vlietland, Rotterdam, the Netherlands (Wils); Department of Intensive Care, Noordwest Ziekenhuisgroep, Alkmaar, the Netherlands (de Ruijter); Department of Intensive Care, Tergooiziekenhuizen, Hilversum, the Netherlands (Vink); Department of Clinical Chemistry, HAGA Ziekenhuis, Den Haag, the Netherlands (Smit).

REFERENCES

1. 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. *Crit Care*. 2008;12(6):R156.

2. Kallet RH, Matthay MA. Hyperoxic acute lung injury. *Respir Care*. 2013;58(1):123-141.

3. Bitterman H, Brod V, Weisz G, Kushnir D, Bitterman N. Effects of oxygen on regional hemodynamics in hemorrhagic shock. *Am J Physiol*. 1996;271(1, pt 2):H203-H211.

4. Cornet AD, Kooter AJ, Peters MJ, Smulders YM. The potential harm of oxygen therapy in medical emergencies. *Crit Care*. 2013;17(2):313.

5. Bone RC, Balk RA, Cerra FB, et al; The ACCP/SCCM Consensus Conference Committee. American College of Chest Physicians/Society of Critical Care Medicine. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Chest*. 1992;101(6):1644-1655.

6. Belda FJ, Aguilera L, García de la Asunción J, et al; Spanish Reduccion de la Tasa de Infeccion Quirurgica Group. Supplemental perioperative oxygen and the risk of surgical wound infection: a randomized controlled trial. *JAMA*. 2005;294(16):2035-2042.

7. Girardis M, Busani S, Damiani E, et al. Effect of conservative vs conventional oxygen therapy on mortality among patients in an intensive care unit. *JAMA*. 2016;316(15):1583-1589.

8. Asfar P, Schortgen F, Boisramé-Helms J, et al; HYPER2S Investigators; REVA research network. Hyperoxia and hypertonic saline in patients with septic shock (HYPERS2S). *Lancet Respir Med*. 2017;5(3):180-190.

9. Panwar R, Hardie M, Bellomo R, et al; CLOSE Study Investigators; ANZICS Clinical Trials Group. Conservative versus liberal oxygenation targets for mechanically ventilated patients. *Am J Respir Crit Care Med*. 2016;193(1):43-51.

10. Mackle D, Bellomo R, Bailey M, et al; ICU-ROX Investigators and the Australian and New Zealand Intensive Care Society Clinical Trials Group; ICU-ROX Investigators 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(11):989-998.

11. Barrot L, Asfar P, Mauny F, et al; LOCO2 Investigators and REVA Research Network. Liberal or conservative oxygen therapy for acute respiratory distress syndrome. *N Engl J Med*. 2020;382(11):999-1008.

12. Schjørring OL, Klitgaard TL, Perner A, et al; HOT-ICU Investigators. Lower or higher oxygenation targets for acute hypoxemic respiratory failure. *N Engl J Med*. 2021;384(14):1301-1311.

13. Chu DK, Kim LH, Young PJ, et al. Mortality and morbidity in acutely ill adults treated with liberal versus conservative oxygen therapy (IOTA). *Lancet*. 2018;391(10131):1693-1705.

14. Ortiz-Prado E, Dunn JF, Vasconez J, Castillo D, Viscor G. Partial pressure of oxygen in the human body. *Am J Blood Res*. 2019;9(1):1-14.

15. Karbing DS, Kjaergaard S, Smith BW, et al. Variation in the PaO2/FiO2 ratio with FiO2. *Crit Care*. 2007;11(6):R118.

16. Whiteley JP, Gavaghan DJ, Hahn CE. Variation of venous admixture, SF6 shunt, PaO2, and the PaO2/FIO2 ratio with FIO2. *Br J Anaesth*. 2002;88(6):771-778.

17. de Grooth HJ, Elbers PWG, Vincent JL. Vitamin C for sepsis and acute respiratory failure. *JAMA*. 2020;323(8):792.

18. de Grooth HJ, Geenen IL, Girbes AR, Vincent JL, Parienti JJ, Oudemans-van Straaten HM. SOFA and mortality endpoints in randomized controlled trials. *Crit Care*. 2017;21(1):38.

19. Hodges Jr JL, Lehmann EL. Comparison of the normal scores and Wilcoxon tests. Paper presented at: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability; 1961.

20. Gordon Lan K, DeMets DL. Discrete sequential boundaries for clinical trials. *Biometrika*. 1983;70(3):659-663.

21. Therneau T. A Package for Survival Analysis in S version 2.38. 2015.

22. Kernan WN, Viscoli CM, Makuch RW, Brass LM, Horwitz RI. Stratified randomization for clinical trials. *J Clin Epidemiol*. 1999;52(1):19-26.

23. Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals. *BMJ*. 2012;345:e5840.

24. Zander R. The oxygen status of arterial human blood. *Scand J Clin Lab Invest Suppl*. 1990;203:187-196.

25. Rice TW, Wheeler AP, Bernard GR, Hayden DL, Schoenfeld DA, Ware LB; National Institutes of Health, National Heart, Lung, and Blood Institute ARDS Network. Comparison of the SpO2/FIO2 ratio and the PaO2/FIO2 ratio in patients with acute lung injury or ARDS. *Chest*. 2007;132(2):410-417.

26. Gershengorn HB, Wunsch H, Scales DC, Zarychanski R, Rubenfeld G, Garland A. Association between arterial catheter use and hospital mortality in intensive care units. *JAMA Intern Med*. 2014;174(11):1746-1754.

27. Smit B, Smulders YM, Eringa EC, et al. Hyperoxia does not affect oxygen delivery in healthy volunteers while causing a decrease in sublingual perfusion. *Microcirculation*. 2018;25(2).

28. Hauser B, Barth E, Bassi G, et al. Hemodynamic, metabolic, and organ function effects of pure oxygen ventilation during established fecal peritonitis-induced septic shock. *Crit Care Med*. 2009;37(8):2465-2469.

29. Barth E, Bassi G, Maybauer DM, et al. Effects of ventilation with 100% oxygen during early hyperdynamic porcine fecal peritonitis. *Crit Care Med*. 2008;36(2):495-503.

30. Vanderweele TJ. Surrogate measures and consistent surrogates. *Biometrics*. 2013;69(3):561-569.

31. de Grooth HJ, Parienti JJ, Oudemans-van Straaten HM. Should we rely on trials with disease- rather than patient-oriented endpoints? *Intensive Care Med*. 2018;44(4):464-466.