

SYSTEMATIC REVIEW



Between-trial heterogeneity in ARDS research

J. Juschten^{1,2,3*} , P. R. Tuinman^{1,2}, T. Guo^{1,2,4}, N. P. Juffermans^{5,6}, M. J. Schultz^{7,8,9}, S. A. Loer³, A. R. J. Girbes^{1,2} and H. J. de Grooth³

© 2021 The Author(s)

Abstract

Purpose: Most randomized controlled trials (RCTs) in patients with acute respiratory distress syndrome (ARDS) revealed indeterminate or conflicting study results. We aimed to systematically evaluate between-trial heterogeneity in reporting standards and trial outcome.

Methods: A systematic review of RCTs published between 2000 and 2019 was performed including adult ARDS patients receiving lung-protective ventilation. A random-effects meta-regression model was applied to quantify heterogeneity (non-random variability) and to evaluate trial and patient characteristics as sources of heterogeneity.

Results: In total, 67 RCTs were included. The 28-day control-group mortality rate ranged from 10 to 67% with large non-random heterogeneity ($I^2 = 88\%$, $p < 0.0001$). Reported baseline patient characteristics explained some of the outcome heterogeneity, but only six trials (9%) reported all four independently predictive variables (mean age, mean lung injury score, mean plateau pressure and mean arterial pH). The 28-day control group mortality adjusted for patient characteristics (i.e. the residual heterogeneity) ranged from 18 to 45%. Trials with significant benefit in the primary outcome reported a higher control group mortality than trials with an indeterminate outcome or harm (mean 28-day control group mortality: 44% vs. 28%; $p = 0.001$).

Conclusion: Among ARDS RCTs in the lung-protective ventilation era, there was large variability in the description of baseline characteristics and significant unexplainable heterogeneity in 28-day control group mortality. These findings signify problems with the generalizability of ARDS research and underline the urgent need for standardized reporting of trial and baseline characteristics.

Keywords: ARDS, Critical Care Research, Heterogeneity

Introduction

The Acute Respiratory Distress Syndrome (ARDS) is a clinically and biologically heterogeneous syndrome that contributes significantly to morbidity and mortality in critically ill patients [1–3]. ARDS has long been a focal point of critical care research, but hundreds of randomized controlled trials (RCTs) have led to merely two

guideline recommendations supported by high-level evidence: low tidal volume ventilation and prone positioning in patients with severe ARDS [4, 5].

The paucity of high-level evidence is due to indeterminate and conflicting trial results. Many RCTs in the ARDS population report an indeterminate outcome—detecting neither significant benefit nor harm of investigated therapeutic strategies [6]. Several other large RCTs demonstrated contradictory results, with seemingly beneficial therapies being found ineffective in subsequent trials [7, 8].

It has become clear that treatment effects of interventions in ARDS are highly dependent on the details of the

*Correspondence: jjuschten@amsterdamumc.nl

¹ Department of Intensive Care, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, mail stop ZH 7D-172, 1081HV Amsterdam, The Netherlands

Full author information is available at the end of the article

intervention, as small variations of the same treatment have led to disparate results. For example, different definitions of ‘low’ and ‘high’ tidal volumes [9–13] or differences in neuromuscular blockade and sedation [7, 8, 14] have led to different trial outcomes. On the other hand, it is much less clear how differences between methodological trial characteristics and patient characteristics affect study outcomes. Between-trial heterogeneity refers to the non-random variation in treatment effect of an intervention due to methodological or clinical differences between patient populations. Unmeasured or unexplainable heterogeneity—both among patients in a single trial and between trial populations—may adversely affect the validity and generalizability of study results (see: ‘Panel: A practical example of the problem with unexplainable between-trial heterogeneity’) [15, 16].

In this study, we set out to quantify the consistency of reporting baseline characteristics and to measure between-trial heterogeneity in 28-day control group mortality among all ARDS RCTs in the lung-protective ventilation era. Our aim was to determine to which extent between-trial differences in control group mortality could be explained by differences in trial and patient characteristics. We hypothesized that between-trial heterogeneity would be large and trial populations often poorly characterized, leading to a discrepancy between inclusion criteria and patient characteristics on the one hand, and control group outcomes on the other hand.

Panel: A practical example of the problem with unexplainable between-trial heterogeneity

We note two high-profile trials published in the same journal issue [17, 18]. Both trials investigated high-frequency oscillatory ventilation in the same target population of moderate to severe ARDS patients, but they reported a different effect on mortality. Judging by the control group patient characteristics, there were clinically meaningful differences between the trial populations: there was a 32% relative difference in baseline Acute Physiology and Chronic Health Evaluation (APACHE II) scores (22 vs. 29 points) and there was a 41% relative difference in control group mortality (41% vs 29% at 30 days). But, paradoxically, the trial with the lowest baseline mean APACHE II score had the highest control group mortality. This makes the interpretation of the conflicting trial results exceedingly difficult. One trial demonstrated significant harm from the intervention while the other trial found no effect. Was the difference in treatment effect due to subtle unreported differences in the intervention, due to unreported differences in the patient populations, or due to differences in the standard of care? Were the patients more severely ill at baseline in the trial with the highest APACHE II score or in the

trial with the highest control group mortality rate? It is clear that unexplainable outcome heterogeneity reduces the generalizability of intervention effects to the global ARDS population [19].

Methods

This systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA). The study protocol and statistical analysis plan were registered online at the International Prospective Register of Systematic Reviews (PROSPERO, registration number: CRD42020161809).

Systematic search

A comprehensive search was conducted in MEDLINE, Embase and Scopus for randomized clinical trials including adult ARDS patients published from January 1st, 2000 until January 31st, 2020. Eligible studies included *a*) adult ARDS patients diagnosed according to the AECC guidelines from 1994 [20] or the Berlin definition from 2012 [21], subjected to *b*) invasive lung-protective mechanical ventilation according to the ARDSnet protocol [11], or reporting a tidal volume of ≤ 8 ml/kg. Included studies were *c*) randomized clinical trials reporting on *d*) 28-day, hospital, intensive care unit (ICU) or 60-day mortality. There were no restrictions with regards to the intervention or phase of the study. More details about the review process are provided in the supplementary appendix, Sect. 1.1 and 1.2.

Outcome measures

For each study, we recorded trial characteristics, intervention, inclusion and exclusion criteria, mean patient baseline characteristics and mortality outcomes.

Primary outcome was the between-trial heterogeneity based on the 28-day control group mortality rate (I^2). The 28-day control group mortality rate reflects the baseline risk of death of a patient population of an individual trial. Secondary outcomes included associations between 28-day control group mortality and characteristics of trial design and outcome, inclusion- and exclusion criteria, as well as baseline characteristics.

Estimation of 28-day control group and intervention group mortality

All analyses investigating heterogeneity were conducted using the 28-day control group mortality rate. For trials reporting solely on the hospital, ICU or 60-day mortality, 28-day control group mortality was estimated with linear regression using data from trials reporting on both, 28-day mortality and any of the other mortality outcomes [22]. 28-day intervention group mortality was estimated in the same manner for analyses investigating differences

between control and intervention group mortality. A sensitivity analysis was conducted, using only the trials reporting 28-day control group mortality.

Estimation and quantification of between-trial heterogeneity

The 28-day control group mortality rates across studies were analyzed using a random-effects meta-regression model with the log odds of mortality as the dependent variable. Each individual trial was weighted by the inverse of the sampling variance of the mortality rates. A maximum likelihood estimator was applied to estimate the mean mortality (random-effects pooled estimate), the between-study standard deviation due to heterogeneity (τ), and heterogeneity (the percentage of variation in control group mortality due to heterogeneity rather than chance, I^2). To make heterogeneity interpretable in a clinically meaningful manner, we calculated the 95% prediction interval. The prediction interval represents the distribution of estimated underlying mortality after correction for random chance and predictive covariates [for further details see method section reference [22]]. This model and its corresponding outcomes were used to present the distribution of 28-day control group mortality between all trials, and to investigate differences between individual trial characteristics.

Associations between patient characteristics and mortality rates

The associations between 28-day control group mortality and reported patient characteristics were estimated by adding each individual covariate separately to the random-effect model as moderators in univariate analysis. The goodness-of-fit of the log-linear, quadratic and power models were compared, and the model with the lowest Akaike information criterion (AIC) was selected [23]. For each model, the regression coefficient (b) and unadjusted R^2 were reported. R^2 represents the proportion of between-trial heterogeneity in 28-day control group mortality explained by the individual baseline characteristic—for the n trials reporting the covariate.

Prediction of control group mortality based on significant patient characteristics

To predict between-trial differences in mortality based on patient characteristics, a comprehensive multivariate logistic regression model was constructed. Missing observations were imputed using multiple imputation generating 20 datasets with predictive mean matching. For a detailed description of the process, we refer to the supplementary appendix, Sect. 2.5. Significant baseline characteristics reported in at least 25% of all trials

with a univariate regression $R^2 \geq 0.10$ were eligible for the model. The threshold R^2 of 0.10 was a compromise between the number of variables and the limited number of observations, as described before [22]. A stepwise backward selection procedure was applied removing regressors if $p \geq 0.05$ for the final model. To facilitate comparisons between the individual covariates the standardized regression coefficient (β) and the standardized standard error (SSE) were reported in the supplementary appendix, Sect. 2.5.

Control group mortality differences between trials demonstrating benefit vs. no benefit

A trial demonstrating significant benefit was defined as a reported p value of < 0.05 for the primary endpoint (as defined by authors) in favor for the intervention group. Comparisons between trials demonstrating significant benefit and trials with an indeterminate outcome or harm were performed using the Mann–Whitney U test. Linear mixed-effects and regression models were applied to estimate the probability of a significant trial outcome based on the observed control group mortality and intervention group mortality, respectively.

Statistical analyses

A p -value of < 0.05 was considered statistically significant.

Statistical analyses were performed with R Studio interface (Version 1.1.447. R core team. R: A Language and Environment for Statistical Computing. 2013. <http://www.r-project.org/>) using the packages ‘tidyverse’, ‘dplyr’, ‘metafor’, ‘mice’, ‘Hmisc’, ‘wCorr’, ‘data.table’, ‘MASS’ and ‘ggplot2’.

Results

Systematic search

The literature search yielded 3479 results. A total of 67 RCTs met all inclusion and exclusion criteria and were included in the analyses (eFigure 1) [7, 8, 17, 18, 24–86]. Table 1 provides an overview of the included trial characteristics.

Estimation of 28-day control group and intervention group mortality

The 28-day control group mortality was reported in 45 trials. For trials reporting another mortality timeframe, 28-day mortality could be reliably estimated (adjusted $R^2 \geq 0.98$ for all estimation models). Linear equations and corresponding regression plots are shown in eTable 1 and eFigure 2. For the sensitivity analysis for trials reporting solely on 28-day mortality we refer to the supplementary appendix, Sect. 2.3.

Table 1 Characteristics of included randomized clinical trials

	Number (%) or median (IQR)
Number of included trials	67
Control group sample size	55 (26 – 169)
Multicenter trials	41 (61)
Trial site*:	
Europe	30 (45)
North America	23 (34)
Middle- and South America	2 (3)
Asia	22 (33)
Australia / New Zealand	6 (9)
Language:	
English	63 (94)
Chinese	4 (6)
Intervention:	
Ventilation strategy	23 (34)
Drug	21 (31)
Prone positioning	4 (9)
Neuromuscular blockage	5 (7)
Nutrition	5 (7)
ECMO	1 (1)
Others	8 (11)
Explicitly stated primary endpoint:	
Mortality	23 (34)
Ventilator-free days at day 28	10 (15)
PaO ₂ /FiO ₂ ratio	9 (13)
Others	18 (26)
Power analysis performed	50 (75)
Early termination of trial	19 (28)
Trials demonstrating significant benefit**	11 (16)
Trials with an indeterminate outcome***	40 (54)
Trials demonstrating significant harm****	4 (6)
Jadad score	2 (1–4)
Composites of the Jadad score:	
Randomization	67 (100)
Randomization adequate	60 (93)
Blinding	28 (42)
Blinding adequate	23 (34)
Description of withdrawals and lost to follow-up	36 (54)

Data are presented as numbers and percentages (%) or median and interquartile range (IQR) according to the type of variable. Legend: *: more than one option possible. **: $p < 0.05$ for the primary outcome (as defined and reported by authors) for the intervention group. ***: $p > 0.05$ for the primary outcome (as defined and reported by authors). ****: $p < 0.05$ for the primary outcome (as defined and reported by authors) for the control group. ECMO extracorporeal membrane oxygenation

Estimation and quantification of between-trial heterogeneity

The 28-day control group mortality ranged from 9.7 to 66.7% with a weighted mean mortality rate of 30.9%.

Between-trial heterogeneity was large with 87.5% of the differences between control group mortality rates not explained by chance ($I^2 = 87.5\%$, $\tau = 0.509$, $p < 0.0001$). The 95% prediction interval (the estimated range of mortality rates, corrected for small trials and random error) was 14 to 55%. The mean mortality and the magnitude of heterogeneity were similar for all subgroups of trial characteristics (Fig. 1). The exclusion criteria were too manifold for valid analyses. However, the number of reported exclusion criteria was associated with a lower 28-day control group mortality ($p = 0.04$, eFigure 3).

Associations between patient characteristics and mortality rates

Table 2 presents the most-reported baseline characteristics and their associations with 28-day control group mortality. For the goodness-of-fit statistics of the individual associations, we refer to eTable 2 in the supplementary appendix. Figure 2 provides an overview of between-trial differences in control group mortality and patient characteristics. We observed an association between 28-day control group mortality and the following patient characteristics in univariate analyses: mean age; mean body mass index; mean APACHE II score; the proportion of patients treated with vasopressors; the proportion of patients presenting with shock at baseline; mean lung injury score; mean oxygenation index; mean plateau pressure; mean PaO₂/FiO₂ ratio and mean arterial pH. Individual regression plots are shown in eFigure 3.

Prediction of control group mortality based on significant patient characteristics

A detailed description of variable selection and construction of the multivariate logistic regression model is available in the supplementary appendix, Sect. 2.5. Significant variables for the final logistic regression model were: mean age ($p < 0.0001$), mean LIS ($p = 0.0099$), mean plateau pressure ($p = 0.0078$) and mean arterial pH ($p = 0.0119$). The residual 95% prediction interval adjusted for the significant predictors was 18 to 45%. Six trials reported all four variables in the final model.

Control group mortality differences between trials demonstrating benefit vs. no benefit

As shown in Fig. 3a, trials demonstrating significant benefit reported a higher 28-day control group mortality compared to trials with an indeterminate outcome or harm (mean 28-day control group mortality rate: 0.275 vs. 0.439; $p = 0.001$). Figure 3b demonstrates that trials with higher control group mortality were more likely to demonstrate significant benefit. Conversely, Fig. 3c shows that intervention group mortality did not differ between trials demonstrating significant benefit and compared

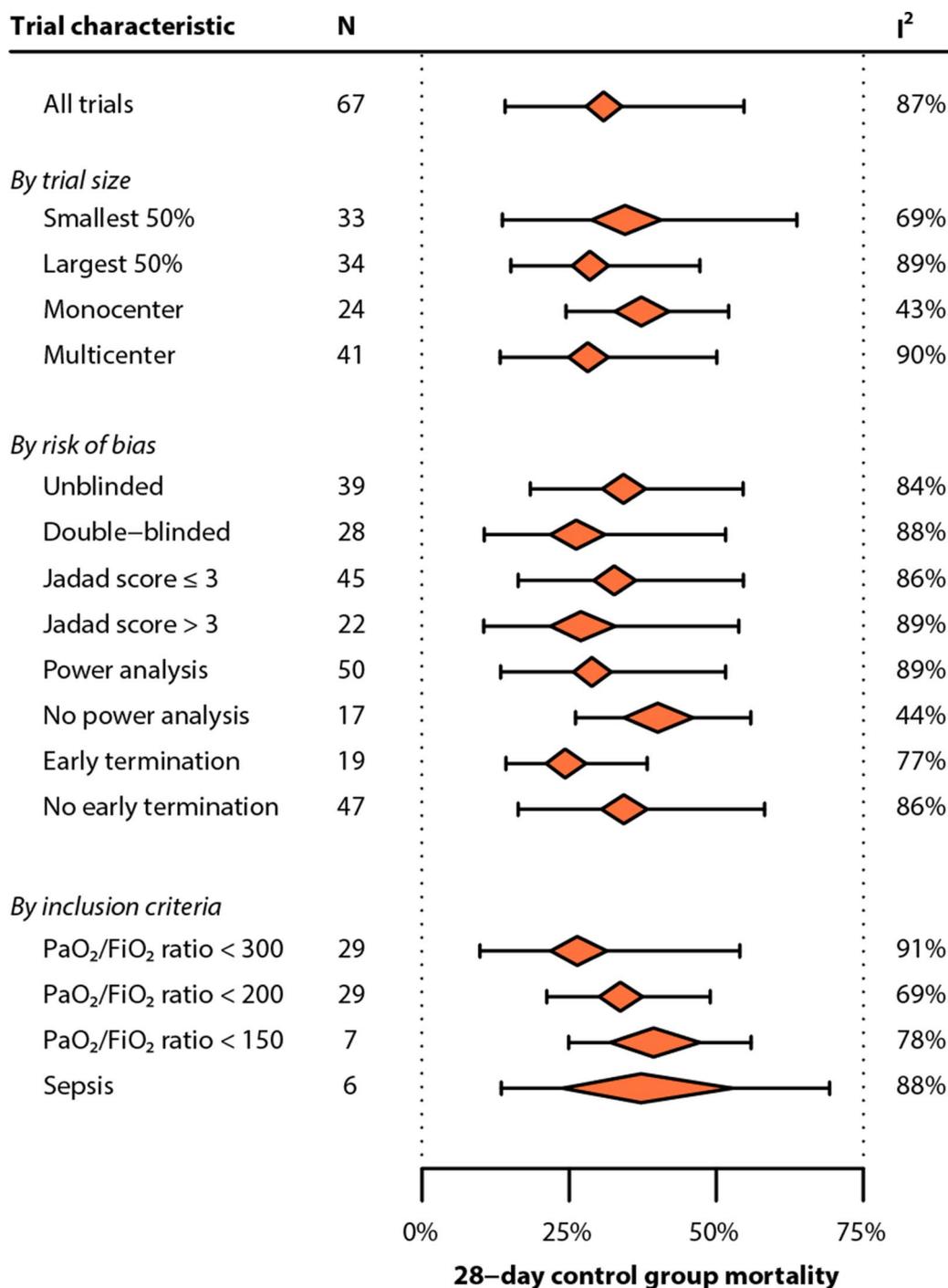


Fig. 1 28-day control group mortality for individual trial characteristics. The diamond represents the mean mortality rate (peak) with the corresponding 95% confidence interval (length of diamond). The black line denotes the 95% prediction interval, which is the estimated between-trial variability in mortality rates after adjusting for random chance and sample size, i.e. the between-trial heterogeneity. I² represents the proportion of between-trial variability that cannot be explained by chance.

to trials with an indeterminate outcome or harm (mean 28-day intervention group mortality rate: 0.271 vs. 0.301; $p = 0.697$). Figure 3d demonstrates that trials with higher

or lower intervention group mortality rates were not more or less likely to demonstrate significant benefit.

Table 2 Univariate associations between 28-day control group mortality and commonly reported mean baseline patient characteristics

Covariate	Number of trials reporting on variable (%)	Mean (SD)	Regression coefficient	R ²	p-value
Publication (year)	67 (100)	2012 (4.6)	− 0.01	0.00	0.429
Age (years)	65 (97)	54.9 (4.3)	0.00	0.16	0.011*
Male gender (%)	60 (90)	59.7 (7.5)	0.76	0.04	0.320
BMI	13 (19)	28.5 (2.3)	− 0.00	0.51	0.004**
Illness severity scores:					
APACHE II	37 (55)	22.6 (3.6)	0.05		0.036*
APACHE III	10 (15)	95 (8.9)	0.01		0.398
APACHE IV	1 (2)	–	–	0.13	–
SAPS II	16 (24)	48.5 (3.3)	0.01	0.14	0.894
SAPS III	1 (2)	–	–	0.00	–
SOFA	24 (36)	9.5 (1.2)	0.07	0.08	0.351
Use of vasopressors (%)	11 (16)	51.1 (17)	0.02	0.058	0.003**
Shock at baseline (%)	4 (6)	56.0 (9.8)	0.00	0.74	0.036*
Risk factors for ARDS (%):					
Pneumonia	44 (66)	51.6 (11.7)	0.01	0.00	0.990
Aspiration	34 (51)	13.1 (5.5)	−0.74	0.00	0.634
Sepsis	31 (46)	31.4 (16.7)	0.13	0.00	0.827
Trauma	24 (37)	6.9 (7.9)	−0.41	0.00	0.746
Transfusion	15 (22)	3.4 (2.9)	6.02	0.08	0.154
Pancreatitis	11 (16)	5.3 (3.2)	0.68	0.00	0.680
Pulmonary severity scores:					
Lung injury score (LIS)	22 (33)	2.7 (0.3)	1.23	0.31	0.023*
Oxygenation index	14 (21)	13.8 (3.2)	0.00	0.35	0.044*
Mechanical ventilation:					
Tidal volume (ml/kg)	50 (75)	7.1 (0.9)	− 0.11	0.05	0.250
Plateau pressure (cmH ₂ O)	45 (67)	25.8 (2.4)	0.09	0.27	0.001**
Minute ventilation (L/min)	21 (31)	10.5 (1.1)	− 0.11	0.10	0.355
Driving pressure (cmH ₂ O)	9 (13)	13.6 (1.9)	0.08	0.16	0.240
PEEP (cm H ₂ O)	51 (76)	10.8 (1.8)	0.04	0.04	0.364
FiO ₂ (%)	21 (31)	0.72 (0.1)	1.95	0.15	0.109
Respiratory rate (/min)	27 (40)	24.6 (2.6)	0.05	0.11	0.120
PaO ₂ /FiO ₂ ratio	60 (90)	134.2 (29.7)	0.00	0.21	0.011*
Arterial pH	26 (39)	7.33 (0.04)	− 5.67	0.50	0.003**
PaCO ₂ (mmHg)	26 (39)	44.9 (4.3)	− 0.00	0.08	0.097

Associations were estimated using a weighted random-effects model with mortality on the log-odds scale. Some baseline characteristics were reported by a minority of trials, which resulted in low power to detect significant associations. R² can be interpreted as the proportion of heterogeneity that is explained by the population characteristic for the *n* trials reporting that characteristic

APACHE Acute Physiology and Chronic Health Evaluation, BMI body mass index, FiO₂ fraction of inspired oxygen, PaCO₂ partial pressure of carbon dioxide in arterial blood, PaO₂ partial pressure of oxygen in arterial blood, SAPS Simplified Acute Physiology Score, SD standard deviation, SOFA Sequential Organ Failure Assessment Score, PEEP positive end-expiratory pressure

Discussion

This systematic analysis of 67 ARDS RCTs in the lung-protective ventilation era revealed a statistically significant and clinically relevant amount of between-trial heterogeneity in reporting and outcome.

The description of patient characteristics was variable and often incomplete. Basic ventilation characteristics

such as mean respiratory rate, FiO₂, pH and PaCO₂ were reported in a minority of trials.

The estimated range of 28-day control group mortality corrected for small trials and random error was 14 to 55%. This between-trial heterogeneity in control group outcomes could not be explained by differences in trial characteristics. Reported baseline patient

(See figure on next page.)

Fig. 2 Heatmap of control group outcomes and baseline characteristics. On the y-axis, all included trials are ordered from highest to lowest (estimated) 28-day mortality rate. The color of a tile represents whether, for a specific trial, a reported variable was lowest (blue) or highest (red) among all trials that reported the variable. A white tile represents a variable not reported by a specific trial. The X-axis depicts the most reported baseline characteristics. Some show a concordant pattern (e.g. age) with 28-day mortality while others do not (e.g. SAPS II score, SOFA score). Most importantly, the distribution of white tiles demonstrates the large variability in the reporting of baseline characteristics.

characteristics explained some of the outcome heterogeneity, but the residual (unexplainable) range in control group mortality was still 18 to 45% after adjusting for the most predictive baseline characteristics. There was no secular trend in mortality outcomes over the period 2000–2019.

Notably, trials with higher control group mortality were more likely to report a significant benefit, also after adjustment for baseline mean severity of illness.

Relevance for clinicians

To assess the applicability of a trial's result to individual patients, clinicians need a clear description of the study population and concomitant treatment from trial reports. In the present study, we identified important problems in this respect.

The variation in reporting of baseline variables was considerable, with $\geq 90\%$ of all studies reporting on age and gender, but only 75% describing observed tidal volumes and PEEP, down to only a third of all trials reporting on lung injury scores or results from blood gas analyses.

After adjustment for significant baseline characteristics (mean age, mean LIS, mean plateau pressure, mean arterial pH), the residual (unexplainable) range in control group mortality was 18 to 45%. In other words, among trials with comparable inclusion criteria and comparable baseline patient characteristics, there were inexplicable 2.5-fold mortality differences (45%/18%). This indicates that there are very important yet unreported differences in ARDS populations, and possibly also differences in co-interventions and standard care.

This silent heterogeneity between trials makes it nearly impossible to evaluate whether RCT results are valid outside of the immediate trial context (i.e., the exact population in the actual participating centers). At its most extreme, this can be thought of as a generalizability crisis in ARDS research: we cannot know which trial results are transportable to which patients outside the trial [87, 88]. The generalizability crisis comes into clear focus when different RCTs show conflicting and statistically mutually exclusive results (benefit vs. no benefit or harm) of the same intervention. Conflicting study results are often ascribed to subtle differences in the intervention, while, in fact, it may be important yet unreported differences in

the population or standard of care that are driving conflicting outcomes.

The finding that the number of exclusion criteria is inversely associated with the control group mortality rate only exacerbates the generalizability problem. It means that trial populations, especially those with a large number of exclusion criteria, likely differ from the intended (broader) target population.

Clinicians trying to gauge the applicability of a trial's result should carefully review not only inclusion criteria and baseline characteristics, but also whether the control group mortality fits the apparent patient characteristics.

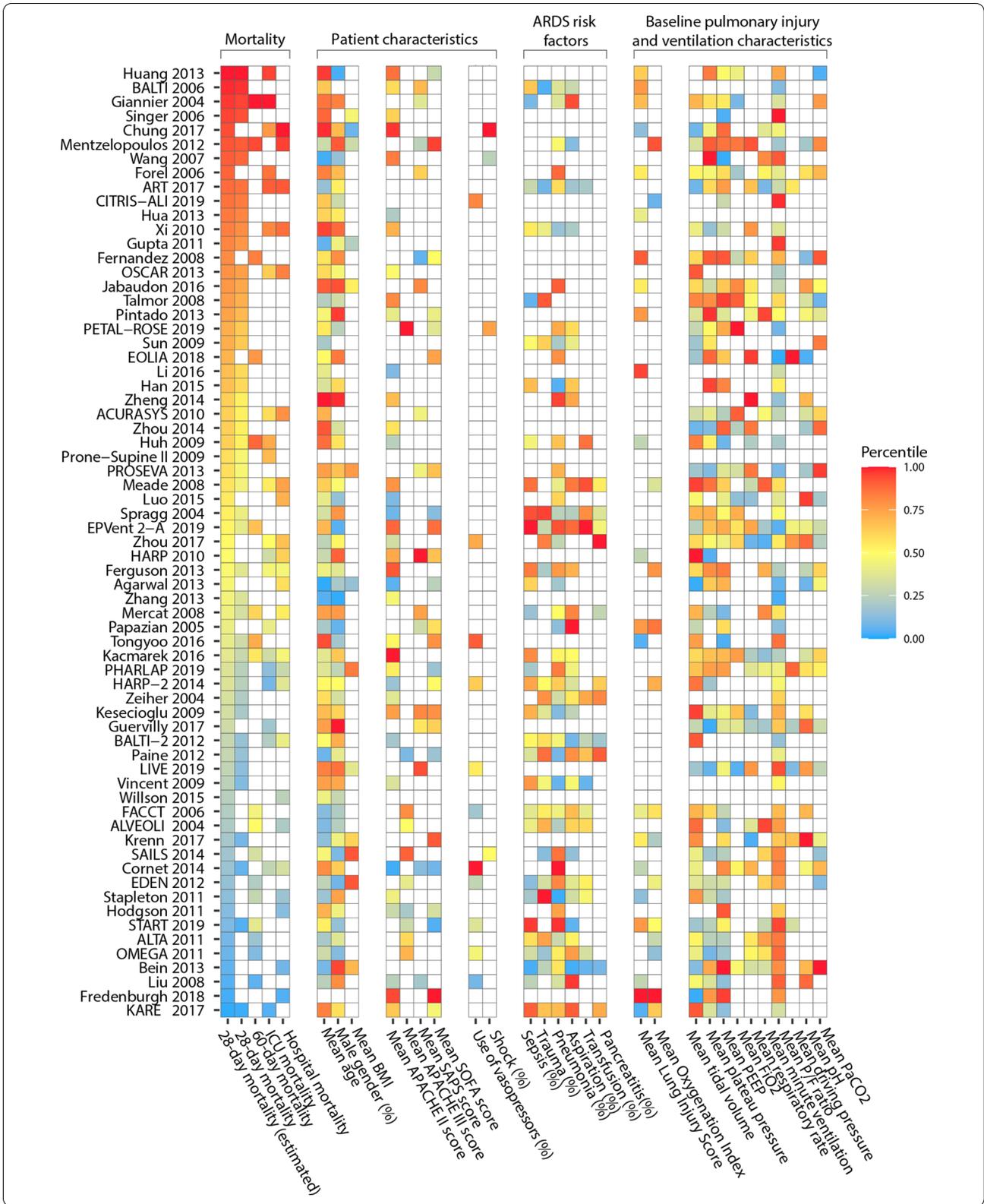
Relevance for ARDS researchers

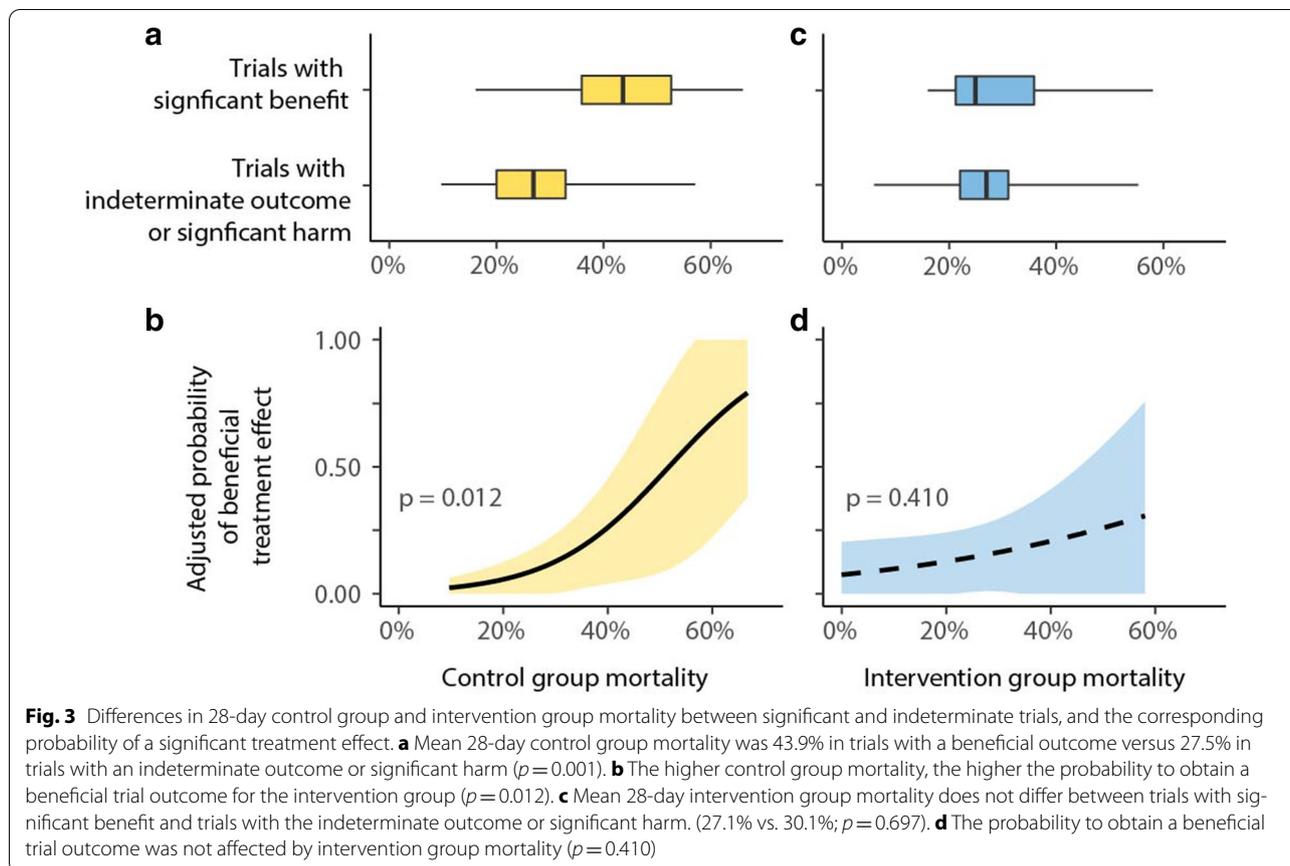
What could account for the large unexplainable heterogeneity? Possible factors may be found in biology, standards of care and co-interventions, or measurement variability.

Statistical cluster analyses of various biological and clinical characteristics led to the identification of distinct ARDS subphenotypes, each associated with a different mortality risk, a different biochemical inflammatory profile and importantly, differential responses to treatments such as PEEP, fluids, low-dose macrolide therapy or simvastatin [89–94]. We are only just beginning to appreciate the extraordinary biological heterogeneity of ARDS [95], which is undoubtedly one of the causes of between-trial variability in outcomes. It is currently not clear whether larger more pragmatic trials or smaller high-adherence trials in more selected populations will provide more useful clinical information in the future.

Variability in standards of care and co-interventions may be another likely cause of between-trial heterogeneity. The LUNG-SAFE study revealed that many ARDS patients are undertreated or not treated according to the current best practices [1]. We cannot know the implications for interpreting RCTs, because standard care and co-interventions are almost never described in trial papers. Historically, this was due to restrictions in space and words allotted by scientific journals. However, there is an urgent need for future studies to report these details, nowadays enabled by the use of supplementary materials and public data repositories.

A final likely cause of between-trial heterogeneity may be measurement variability of baseline characteristics and clinical outcomes. Severity of illness scores, for





example, are notoriously dependent on small variations in measurement definitions [96, 97]. The call for standardization of baseline and outcome measures in ARDS trials is not new, but the current study underlines its urgency and importance [98, 99].

Consequently, comprehensive reporting in a standardized manner on patient characteristics, standard care and concomitant treatments may be one step towards solving the generalizability crisis in ARDS research. Clearly, heterogeneity is not the only cause of a large number of indeterminate and conflicting trial results. Statistical shortcomings, such as underpowered studies [100] or overestimations of effect size [101], equally contribute to indeterminate trial outcomes. Moreover, qualifying studies as ‘indeterminate’ is the consequence of the frequentist statistical paradigm, while Bayesian analyses offer another perspective providing often useful information about trial results [102].

Because we were limited to trial-level data in this study, we should be careful to avoid the ‘ecological fallacy’: individual-level relationships cannot be inferred from group-level data. This important limitation of the present study is also an important message: we will

continue to fail to understand outcome heterogeneity between ARDS trials as long as we must rely on aggregated study-level data. Sharing (anonymized) individual patient data is likely to provide a path forward and provides an immense opportunity to stratify and subphenotype patients, to detect treatment benefit and harm for specific patient groups, and to find valuable therapeutic strategies in an inherently heterogeneous syndrome.

Complete standardization of reporting is unwarranted and can even be detrimental. Reported characteristics and outcomes should be tailored to the research question at hand. But the results in this study indicate that different trials lack sufficient common ground to validly compare trial populations. Creating this common ground for between-trial comparisons requires the reporting of a ‘core baseline set’: a commonly agreed-upon minimum set of descriptors to characterize the patient population of a trial (including standards of care and co-interventions). Developing this core population characteristics set requires meta-epidemiological data (which this study provides) and clinical domain expertise from a diverse sample of ARDS researchers.

Conclusion

Randomized controlled ARDS trials in the lung-protective ventilation era present a statistically significant and clinically relevant amount of heterogeneity in reporting and mortality outcomes. Differences in baseline characteristics partly explained the variability in outcome, but large unexplainable heterogeneity remained after extensive statistical adjustments. This study underlines the urgent need for standardized and comprehensive reporting of trial and baseline characteristics to diminish between-trial heterogeneity and to support the transportability of study results across populations.

Supplementary Information

The online version of this article (<https://doi.org/10.1007/s00134-021-06370-w>) contains supplementary material, which is available to authorized users.

Author details

¹ Department of Intensive Care, Amsterdam UMC, Vrije Universiteit Amsterdam, De Boelelaan 1117, mail stop ZH 7D-172, 1081HV Amsterdam, The Netherlands. ² Research VUmc Intensive Care (REVIVE), Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. ³ Department of Anesthesiology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. ⁴ Division of System Biomedicine and Pharmacology, Leiden Academic Centre for Drug Research (LACDR), Leiden University, Leiden, The Netherlands. ⁵ Laboratory of Experimental Intensive Care and Anesthesiology (LEICA), Amsterdam UMC, Universiteit Van Amsterdam, Amsterdam, The Netherlands. ⁶ Department of Intensive Care, OLVG Hospital, Amsterdam, The Netherlands. ⁷ Department of Intensive Care, Amsterdam UMC, Universiteit Van Amsterdam, Amsterdam, The Netherlands. ⁸ Mahidol–Oxford Tropical Medicine Research Unit (MORU), Mahidol University, Bangkok, Thailand. ⁹ Nuffield Department of Medicine, University of Oxford, Oxford, UK.

Authors' contributions

JJ, PRT, NPJ, MJS, SAL, ARJG and HJdG developed the study design of the present study. JJ, TG and HJdG conducted the search and retrieved trial information according to the PRISMA guidelines. JJ and HJdG performed the statistical analyses. JJ, PRT, TG, NPJ, MJS, SAL, ARJG and HJdG interpreted the data. JJ, PRT, NPJ, MJS, ARJ and HJdG drafted the manuscript. All authors reviewed the manuscript and approved the final submitted version.

Declarations

Competing interests

The authors declare that there is no conflict of interest.

Open Access

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Published online: 13 March 2021

References

- Bellani G, Laffey JG, Pham T, Fan E, Brochard L, Esteban A et al (2016) Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 Countries. *JAMA* 315(8):788–800
- Rubinfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M et al (2005) Incidence and outcomes of acute lung injury. *N Engl J Med* 353(16):1685–1693
- Villar J, Blanco J, Añón JM, Santos-Bouza A, Blanch L, Ambrós A et al (2011) The ALIEN study: incidence and outcome of acute respiratory distress syndrome in the era of lung protective ventilation. *Intensive Care Med* 37(12):1932–1941
- Fan E, Del Sorbo L, Goligher EC, Hodgson CL, Munshi L, Walkey AJ et al (2017) An Official American Thoracic Society/European Society of Intensive Care Medicine/Society Of Critical Care Medicine clinical practice guideline: mechanical ventilation in adult patients with acute respiratory distress syndrome. *Am J Respir Crit Care Med* 195(9):1253–1263
- Papazian L, Aubron C, Brochard L, Chiche JD, Combes A, Dreyfuss D et al (2019) Formal guidelines: management of acute respiratory distress syndrome. *Ann Intensive Care* 9(1):69
- Tonelli AR, Zein J, Adams J, Ioannidis JP (2014) Effects of interventions on survival in acute respiratory distress syndrome: an umbrella review of 159 published randomized trials and 29 meta-analyses. *Intensive Care Med* 40(6):769–787
- National Heart L, Blood Institute PCTN, Moss M, Huang DT, Brower RG, Ferguson ND et al (2019) Early neuromuscular blockade in the acute respiratory distress syndrome. *N Engl J Med* 380(21):1997–2008
- Papazian L, Forel JM, Gacouin A, Penot-Ragon C, Perrin G, Loundou A et al (2010) Neuromuscular blockers in early acute respiratory distress syndrome. *N Engl J Med* 363(12):1107–1116
- Amato MB, Barbas CS, Medeiros DM, Magaldi RB, Schettino GP, Lorenzi-Filho G et al (1998) Effect of a protective-ventilation strategy on mortality in the acute respiratory distress syndrome. *N Engl J Med* 338(6):347–354
- Brochard L, Roudot-Thoraval F, Roupie E, Delclaux C, Chastre J, Fernandez-Mondejar E et al (1998) Tidal volume reduction for prevention of ventilator-induced lung injury in acute respiratory distress syndrome. The Multicenter Trial Group on Tidal Volume reduction in ARDS. *Am J Respir Crit Care Med* 158(6):1831–1838
- Brower RG, Matthay MA, Morris A, Schoenfeld D, Thompson BT, Wheeler A (2000) Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. *N Engl J Med* 342(18):1301–1308
- Brower RG, Shanholtz CB, Fessler HE, Shade DM, White P Jr, Wiener CM et al (1999) Prospective, randomized, controlled clinical trial comparing traditional versus reduced tidal volume ventilation in acute respiratory distress syndrome patients. *Crit Care Med* 27(8):1492–1498
- Stewart TE, Meade MO, Cook DJ, Granton JT, Hodder RV, Lapinsky SE et al (1998) Evaluation of a ventilation strategy to prevent barotrauma in patients at high risk for acute respiratory distress syndrome: Pressure- and Volume-Limited Ventilation Strategy Group. *N Engl J Med* 338(6):355–361
- Tarazan N, Alshehri M, Sharif S, Al Duhaib Z, Moller MH, Belley-Cote E et al (2020) Neuromuscular blocking agents in acute respiratory distress syndrome: updated systematic review and meta-analysis of randomized trials. *Intensive Care Med* 8(1):61
- Iwashyna TJ, Burke JF, Sussman JB, Prescott HC, Hayward RA, Angus DC (2015) Implications of heterogeneity of treatment effect for reporting and analysis of randomized trials in critical care. *Am J Respir Crit Care Med* 192(9):1045–1051
- Santhakumaran S, Gordon A, Prevost AT, O'Kane C, McAuley DF, Shankar-Hari M (2019) Heterogeneity of treatment effect by baseline risk of mortality in critically ill patients: re-analysis of three recent sepsis and ARDS randomised controlled trials. *Crit Care* 23(1):156

17. Ferguson ND, Cook DJ, Guyatt GH, Mehta S, Hand L, Austin P et al (2013) High-frequency oscillation in early acute respiratory distress syndrome. *N Engl J Med* 368(9):795–805
18. Young D, Lamb SE, Shah S, MacKenzie I, Tunnicliffe W, Lall R et al (2013) High-frequency oscillation for acute respiratory distress syndrome. *N Engl J Med* 368(9):806–813
19. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR (2017) Generalizing study results: a potential outcomes perspective. *Epidemiology* 28(4):553–561
20. Bernard GR, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L et al (1994) The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 149(3 Pt 1):818–824
21. Force ADT, Ranieri VM, Rubenfeld GD, Thompson BT, Ferguson ND, Caldwell E et al (2012) Acute respiratory distress syndrome: the Berlin Definition. *JAMA* 307(23):2526–2533
22. de Grooth HJ, Postema J, Loer SA, Parienti JJ, Oudemans-van Straaten HM, Girbes AR (2018) Unexplained mortality differences between septic shock trials: a systematic analysis of population characteristics and control-group mortality rates. *Intensive Care Med* 44(3):311–322
23. Moses T, Holland PW (2010) A comparison of statistical selection strategies for univariate and bivariate log-linear models. *Br J Math Stat Psychol* 63(Pt 3):557–574
24. Agarwal R, Srinivasan A, Aggarwal AN, Gupta D (2013) Adaptive support ventilation for complete ventilatory support in acute respiratory distress syndrome: a pilot, randomized controlled trial. *Respirology* 18(7):1108–1115
25. Bein T, Weber-Carstens S, Goldmann A, Muller T, Staudinger T, Brederlau J et al (2013) Lower tidal volume strategy (approximately 3 ml/kg) combined with extracorporeal CO₂ removal versus “conventional” protective ventilation (6 ml/kg) in severe ARDS: the prospective randomized Xtravent-study. *Intensive Care Med* 39(5):847–856
26. Beitler JR, Sarge T, Banner-Goodspeed VM, Gong MN, Cook D, Novack V et al (2019) Effect of titrating positive end-expiratory pressure (PEEP) with an esophageal pressure-guided strategy vs an empirical high PEEP-FiO₂ strategy on death and days free from mechanical ventilation among patients with acute respiratory distress syndrome: a randomized clinical trial. *JAMA* 321(9):846–857
27. Brower RG, Lanken PN, MacIntyre N, Matthay MA, Morris A, Ancukiewicz M et al (2004) Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. *N Engl J Med* 351(4):327–336
28. Cavalcanti AB, Suzumura EA, Laranjeira LN, Paisani DM, Damiani LP, Guimaraes HP et al (2017) Effect of lung recruitment and titrated positive end-expiratory pressure (PEEP) vs low PEEP on mortality in patients with acute respiratory distress syndrome: a randomized clinical trial. *JAMA* 318(14):1335–1345
29. Chung FT, Lee CS, Lin SM, Kuo CH, Wang TY, Fang YF et al (2017) Alveolar recruitment maneuver attenuates extravascular lung water in acute respiratory distress syndrome. *Medicine (Baltimore)* 96(30):e7627
30. Combes A, Hajage D, Capellier G, Demoule A, Lavoue S, Guervilly C et al (2018) Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. *N Engl J Med* 378(21):1965–1975
31. Constantin JM, Jabaudon M, Lefrant JY, Jaber S, Quenot JP, Langeron O et al (2019) Personalised mechanical ventilation tailored to lung morphology versus low positive end-expiratory pressure for patients with acute respiratory distress syndrome in France (the LIVE study): a multicentre, single-blind, randomised controlled trial. *The Lancet Respiratory medicine* 7(10):870–880
32. Cornet AD, Groeneveld AB, Hofstra JJ, Vlaar AP, Tuinman PR, van Lingen A et al (2014) Recombinant human activated protein C in the treatment of acute respiratory distress syndrome: a randomized clinical trial. *PLoS ONE* 9(3):e90983
33. Craig TR, Duffy MJ, Shyamsundar M, McDowell C, O’Kane CM, Elborn JS et al (2011) A randomized clinical trial of hydroxymethylglutaryl-coenzyme A reductase inhibition for acute lung injury (The HARP Study). *Am J Respir Crit Care Med* 183(5):620–626
34. Fernandez R, Trenches X, Klamburg J, Castedo J, Serrano JM, Besso G et al (2008) Prone positioning in acute respiratory distress syndrome: a multicenter randomized clinical trial. *Intensive Care Med* 34(8):1487–1491
35. Forel JM, Roch A, Marin V, Michelet P, Demory D, Blache JL et al (2006) Neuromuscular blocking agents decrease inflammatory response in patients presenting with acute respiratory distress syndrome. *Crit Care Med* 34(11):2749–2757
36. Fowler AA, Truitt JD, Hite RD, Morris PE, DeWilde C, Priday A et al (2019) Effect of Vitamin C Infusion on Organ Failure and Biomarkers of inflammation and vascular injury in patients with sepsis and severe acute respiratory failure: the CITRIS-ALI randomized clinical trial. *JAMA* 322(13):1261–1270
37. Fredenburgh LE, Perrella MA, Barragan-Bradford D, Hess DR, Peters E, Welty-Wolf KE et al (2018) A phase I trial of low-dose inhaled carbon monoxide in sepsis-induced ARDS. *JCI Insight*. 3(23):124039
38. Gannier M, Roch A, Forel JM, Thirion X, Arnal JM, Donati S et al (2004) Effect of neuromuscular blocking agents on gas exchange in patients presenting with acute respiratory distress syndrome. *Crit Care Med* 32(1):113–119
39. Gao Smith F, Perkins GD, Gates S, Young D, McAuley DF, Tunnicliffe W et al (2012) Effect of intravenous beta-2 agonist treatment on clinical outcomes in acute respiratory distress syndrome (BALTI-2): a multicentre, randomised controlled trial. *Lancet* 379(9812):229–235
40. Guerin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T et al (2013) Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med* 368(23):2159–2168
41. Guervilly C, Bisbal M, Forel JM, Mechat M, Lehingue S, Bourenne J et al (2017) Effects of neuromuscular blockers on transpulmonary pressures in moderate to severe acute respiratory distress syndrome. *Intensive Care Med* 43(3):408–418
42. Gupta A, Govil D, Bhatnagar S, Gupta S, Goyal J, Patel S et al (2011) Efficacy and safety of parenteral omega 3 fatty acids in ventilated patients with acute lung injury. *Indian J Crit Care Med* 15(2):108–113
43. Han F, Sun R, Ni Y, Hu X, Chen X, Jiang L et al (2015) Early initiation of continuous renal replacement therapy improves clinical outcomes in patients with acute respiratory distress syndrome. *Am J Med Sci* 349(3):199–205
44. Hodgson CL, Cooper DJ, Arabi Y, King V, Bersten A, Bihari S et al (2019) Maximal recruitment open lung ventilation in acute respiratory distress syndrome (PHARLAP). A phase II, multicenter randomized controlled clinical trial. *Am J Respir Crit Care Med*. 200(11):1363–1372
45. Hodgson CL, Tuxen DV, Davies AR, Bailey MJ, Higgins AM, Holland AE et al (2011) A randomised controlled trial of an open lung strategy with staircase recruitment, titrated PEEP and targeted low airway pressures in patients with acute respiratory distress syndrome. *Crit Care* 15(3):R133
46. Hua F, Wang X, Zhu L (2013) Terlipressin decreases vascular endothelial growth factor expression and improves oxygenation in patients with acute respiratory distress syndrome and shock. *J Emerg Med* 44(2):434–439
47. Huang Z, Wang SR, Yang ZL, Liu JY (2013) Effect on extrapulmonary sepsis-induced acute lung injury by hemoperfusion with neutral microporous resin column. *Ther Apheresis Dial* 17(4):454–461
48. Huh JW, Jung H, Choi HS, Hong SB, Lim CM, Koh Y (2009) Efficacy of positive end-expiratory pressure titration after the alveolar recruitment manoeuvre in patients with acute respiratory distress syndrome. *Crit Care* 13(1):R22
49. Jabaudon M, Boucher P, Imhoff E, Chabanne R, Faure JS, Roszyk L et al (2017) Sevoflurane for sedation in acute respiratory distress syndrome. A randomized controlled pilot study. *Am J Respir Crit Care Med*. 195(6):792–800
50. Kacmarek RM, Villar J, Sulemanji D, Montiel R, Ferrando C, Blanco J et al (2016) Open lung approach for the acute respiratory distress syndrome: a pilot, Randomized Controlled Trial. *Crit Care Med* 44(1):32–42
51. Kesecioglu J, Beale R, Stewart TE, Findlay GP, Rouby JJ, Holzapfel L et al (2009) Exogenous natural surfactant for treatment of acute lung injury and the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 180(10):989–994
52. Krenn K, Lucas R, Croize A, Boehme S, Klein KU, Hermann R et al (2017) Inhaled AP301 for treatment of pulmonary edema in mechanically ventilated patients with acute respiratory distress syndrome: a phase IIa randomized placebo-controlled trial. *Crit Care* 21(1):194
53. Li JQ, Li N, Han GJ, Pan CG, Zhang YH, Shi XZ et al (2016) Clinical research about airway pressure release ventilation for moderate to

- severe acute respiratory distress syndrome. *Eur Rev Med Pharmacol Sci* 20(12):2634–2641
54. Liu KD, Levitt J, Zhuo H, Kallet RH, Brady S, Steingrub J et al (2008) Randomized clinical trial of activated protein C for the treatment of acute lung injury. *Am J Respir Crit Care Med* 178(6):618–623
 55. Luo J, Wang MY, Liang BM, Yu H, Jiang FM, Wang T et al (2015) Initial synchronized intermittent mandatory ventilation versus assist/control ventilation in treatment of moderate acute respiratory distress syndrome: a prospective randomized controlled trial. *J Thorac Dis* 7(12):2262–2273
 56. Matthay MA, Brower RG, Carson S, Douglas IS, Eisner M, Hite D et al (2011) Randomized, placebo-controlled clinical trial of an aerosolized β_2 -agonist for treatment of acute lung injury. *Am J Respir Crit Care Med* 184(5):561–568
 57. Matthay MA, Calfee CS, Zhuo H, Thompson BT, Wilson JG, Levitt JE et al (2019) Treatment with allogeneic mesenchymal stromal cells for moderate to severe acute respiratory distress syndrome (START study): a randomised phase 2a safety trial. *The Lancet Respiratory medicine* 7(2):154–162
 58. McAuley DF, Cross LM, Hamid U, Gardner E, Elborn JS, Cullen KM et al (2017) Keratinocyte growth factor for the treatment of the acute respiratory distress syndrome (KARE): a randomised, double-blind, placebo-controlled phase 2 trial. *Lancet Respir Med* 5(6):484–491
 59. McAuley DF, Laffey JG, O’Kane CM, Perkins GD, Mullan B, Trinder TJ et al (2014) Simvastatin in the acute respiratory distress syndrome. *N Engl J Med* 371(18):1695–1703
 60. Meade MO, Cook DJ, Guyatt GH, Slutsky AS, Arabi YM, Cooper DJ et al (2008) Ventilation strategy using low tidal volumes, recruitment maneuvers, and high positive end-expiratory pressure for acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 299(6):637–645
 61. Mentzelopoulos SD, Malachias S, Zintzaras E, Kokkoris S, Zakyntinos E, Makris D et al (2012) Intermittent recruitment with high-frequency oscillation/tracheal gas insufflation in acute respiratory distress syndrome. *Eur Respir J* 39(3):635–647
 62. Mercat A, Richard JC, Vieille B, Jaber S, Osman D, Diehl JL et al (2008) Positive end-expiratory pressure setting in adults with acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 299(6):646–655
 63. Paine R 3rd, Standiford TJ, Dechert RE, Moss M, Martin GS, Rosenberg AL et al (2012) A randomized trial of recombinant human granulocyte-macrophage colony stimulating factor for patients with acute lung injury. *Crit Care Med* 40(1):90–97
 64. Papazian L, Gannier M, Marin V, Donati S, Arnal JM, Demory D et al (2005) Comparison of prone positioning and high-frequency oscillatory ventilation in patients with acute respiratory distress syndrome. *Crit Care Med* 33(10):2162–2171
 65. Perkins GD, McAuley DF, Thickett DR, Gao F (2006) The beta-agonist lung injury trial (BALTI): a randomized placebo-controlled clinical trial. *Am J Respir Crit Care Med* 173(3):281–287
 66. Pintado MC, de Pablo R, Trascasa M, Milicua JM, Rogero S, Daguerre M et al (2013) Individualized PEEP setting in subjects with ARDS: a randomized controlled pilot study. *Respir Care* 58(9):1416–1423
 67. Rice TW, Wheeler AP, Thompson BT, deBoisblanc BP, Steingrub J, Rock P (2011) Enteral omega-3 fatty acid, gamma-linolenic acid, and antioxidant supplementation in acute lung injury. *JAMA* 306(14):1574–1581
 68. Rice TW, Wheeler AP, Thompson BT, Steingrub J, Hite RD, Moss M et al (2012) Initial trophic vs full enteral feeding in patients with acute lung injury: the EDEN randomized trial. *JAMA* 307(8):795–803
 69. Singer P, Theilla M, Fisher H, Gibstein L, Grozovski E, Cohen J (2006) Benefit of an enteral diet enriched with eicosapentaenoic acid and gamma-linolenic acid in ventilated patients with acute lung injury. *Crit Care Med* 34(4):1033–1038
 70. Spragg RG, Lewis JF, Walrath HD, Johannigman J, Bellingan G, Laterre PF et al (2004) Effect of recombinant surfactant protein C-based surfactant on the acute respiratory distress syndrome. *N Engl J Med* 351(9):884–892
 71. Stapleton RD, Martin TR, Weiss NS, Crowley JJ, Gundel SJ, Nathens AB et al (2011) A phase II randomized placebo-controlled trial of omega-3 fatty acids for the treatment of acute lung injury. *Crit Care Med* 39(7):1655–1662
 72. Sun JJ, Yang MW, Wang CH, Li CH, Qin WH, Li YL et al (2009) Clinical effects of low-stretch ventilation on acute respiratory distress syndrome. *Chin Crit Care Med* 21(10):609–612
 73. Taccone P, Pesenti A, Latini R, Polli F, Vagginelli F, Mietto C et al (2009) Prone positioning in patients with moderate and severe acute respiratory distress syndrome: a randomized controlled trial. *JAMA* 302(18):1977–1984
 74. Talmor D, Sarge T, Malhotra A, O’Donnell CR, Ritz R, Lisbon A et al (2008) Mechanical ventilation guided by esophageal pressure in acute lung injury. *N Engl J Med* 359(20):2095–2104
 75. Tongyoo S, Permpikul C, Mongkolpun W, Vattanavanit V, Udompanturak S, Kocak M et al (2016) Hydrocortisone treatment in early sepsis-associated acute respiratory distress syndrome: results of a randomized controlled trial. *Crit Care* 20(1):329
 76. Truwit JD, Bernard GR, Steingrub J, Matthay MA, Liu KD, Albertson TE et al (2014) Rosuvastatin for sepsis-associated acute respiratory distress syndrome. *N Engl J Med* 370(23):2191–2200
 77. Vincent JL, Artigas A, Petersen LC, Meyer C (2009) A multicenter, randomized, double-blind, placebo-controlled, dose-escalation trial assessing safety and efficacy of active site inactivated recombinant factor VIIa in subjects with acute lung injury or acute respiratory distress syndrome. *Crit Care Med* 37(6):1874–1880
 78. Wang XZ, Lu CJ, Gao FQ, Li XH, Hao D, Ning FY (2007) Comparison of the effects of BiPAP ventilation combined with lung recruitment maneuvers and low tidal volume A/C ventilation in patients with acute respiratory distress syndrome. *Chin J Tuberculosis Respir Dis* 30(1):44–47
 79. Wiedemann HP, Wheeler AP, Bernard GR, Thompson BT, Hayden D, deBoisblanc B et al (2006) Comparison of two fluid-management strategies in acute lung injury. *N Engl J Med* 354(24):2564–2575
 80. Willson DF, Truwit JD, Conaway MR, Traul CS, Egan EE (2015) The Adult Calfactant in Acute Respiratory Distress Syndrome Trial. *Chest* 148(2):356–364
 81. Xi XM, Jiang L, Zhu B (2010) Clinical efficacy and safety of recruitment maneuver in patients with acute respiratory distress syndrome using low tidal volume ventilation: a multicenter randomized controlled clinical trial. *Chin Med J* 123(21):3100–3105
 82. Zeiher BG, Artigas A, Vincent JL, Dmitrienko A, Jackson K, Thompson BT et al (2004) Neutrophil elastase inhibition in acute lung injury: results of the STRIVE study. *Crit Care Med* 32(8):1695–1702
 83. Zhang JC, Chu YF, Zeng J, Ren HS, Meng M, Jiang JJ et al (2013) Effect of continuous high-volume hemofiltration in patients with severe acute respiratory distress syndrome. *Zhonghua wei zhong bing ji jiu yi xue* 25(3):145–148
 84. Zheng G, Huang L, Tong H, Shu Q, Hu Y, Ge M et al (2014) Treatment of acute respiratory distress syndrome with allogeneic adipose-derived mesenchymal stem cells: a randomized, placebo-controlled pilot study. *Respir Res* 15:39
 85. Zhou X, Liu D, Long Y, Zhang Q, Cui N, He H et al (2014) The effects of prone position ventilation combined with recruitment maneuvers on outcomes in patients with severe acute respiratory distress syndrome. *Zhonghua Nei Ke Za Zhi* 53(6):437–441
 86. Zhou Y, Jin X, Lv Y, Wang P, Yang Y, Liang G et al (2017) Early application of airway pressure release ventilation may reduce the duration of mechanical ventilation in acute respiratory distress syndrome. *Intensive Care Med* 43(11):1648–1659
 87. Hernán MA, VanderWeele TJ (2011) Compound treatments and transportability of causal inference. *Epidemiology* 22(3):368–377
 88. Westreich D, Edwards JK, Lesko CR, Stuart E, Cole SR (2017) Transportability of Trial Results Using Inverse Odds of Sampling Weights. *Am J Epidemiol* 186(8):1010–1014
 89. Bos LD, Schouten LR, van Vught LA, Wiewel MA, Ong DSY, Cremer O et al (2017) Identification and validation of distinct biological phenotypes in patients with acute respiratory distress syndrome by cluster analysis. *Thorax* 72(10):876–883
 90. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA et al (2014) Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. *Lancet Respir Med* 2(8):611–620
 91. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M et al (2018) Acute respiratory distress syndrome subphenotypes and

- differential response to simvastatin: secondary analysis of a randomised controlled trial. *Lancet Respir Med* 6(9):691–698
92. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT et al (2017) Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. *Am J Respir Crit Care Med* 195(3):331–338
 93. Simonis FD, de Iudicibus G, Cremer OL, Ong DSY, van der Poll T, Bos LD et al (2018) Macrolide therapy is associated with reduced mortality in acute respiratory distress syndrome (ARDS) patients. *Ann Transl Med* 6(2):24
 94. Sinha P, Delucchi KL, Thompson BT, McAuley DF, Matthay MA, Calfee CS et al (2018) Latent class analysis of ARDS subphenotypes: a secondary analysis of the statins for acutely injured lungs from sepsis (SAILS) study. *Intensive Care Med* 44(11):1859–1869
 95. Bos LDJ, Scicluna BP, Ong DSY, Cremer O, van der Poll T, Schultz MJ (2019) Understanding heterogeneity in biologic phenotypes of acute respiratory distress syndrome by leukocyte expression profiles. *Am J Respir Crit Care Med* 200(1):42–50
 96. Féry-Lemonnier E, Landais P, Loirat P, Kleinknecht D, Brivet F (1995) Evaluation of severity scoring systems in ICUs—translation, conversion and definition ambiguities as a source of inter-observer variability in Apache II. *SAPS and OSF Intens Care Med* 21(4):356–360
 97. Klein Klouwenberg PM, Ong DS, Bonten MJ, Cremer OL (2012) Classification of sepsis, severe sepsis and septic shock: the impact of minor variations in data capture and definition of SIRS criteria. *Intensive Care Med* 38(5):811–819
 98. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO et al (2017) Core outcome measures for clinical research in acute respiratory failure survivors an international modified delphi consensus study. *Am J Respir Crit Care Med*. 196(9):1122–1130
 99. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST et al (2017) The COMET Handbook: version 1.0. *Trials* 18(Suppl 3):280
 100. Rubenfeld GD (2015) Confronting the frustrations of negative clinical trials in acute respiratory distress syndrome. *Ann Am Thoracic Soc* 12(Suppl 1):S58–63
 101. Aberlegg SK, Richards DR, O'Brien JM (2010) Delta inflation: a bias in the design of randomized controlled trials in critical care medicine. *Crit Care* 14(2):R77
 102. Yarnell CJ, Abrams D, Baldwin MR, Brodie D, Fan E, Ferguson ND et al (2021) Clinical trials in critical care: can a Bayesian approach enhance clinical and scientific decision making? *Lancet Respir Med* 9(2):207–216