Intensive Care Med (2020) 46:930–942
https://doi.org/10.1007/s00134-020-05934-6

CONFERENCE REPORTS AND EXPERT PANEL

Contemporary strategies to improve clinical trial design for critical care research: insights from the First Critical Care Clinical Trialists Workshop

Michael O. Harhay[1,2*], Jonathan D. Casey[3], Marina Clement[4], Sean P. Collins[5], Étienne Gayat[6,7,8], Michelle Ng Gong[9], Samir Jaber[10], Pierre-François Laterre[11], John C. Marshall[12], Michael A. Matthay[13], Rhonda E. Monroe[14], Todd W. Rice[3], Eileen Rubin[15], Wesley H. Self[5] and Alexandre Mebazaa[6,7,8]

Abstract

Background: Conducting research in critically-ill patient populations is challenging, and most randomized trials of critically-ill patients have not achieved pre-specified statistical thresholds to conclude that the intervention being investigated was beneficial.

Methods: In 2019, a diverse group of patient representatives, regulators from the USA and European Union, federal grant managers, industry representatives, clinical trialists, epidemiologists, and clinicians convened the First Critical Care Clinical Trialists (3CT) Workshop to discuss challenges and opportunities in conducting and assessing critical care trials. Herein, we present the advantages and disadvantages of available methodologies for clinical trial design, conduct, and analysis, and a series of recommendations to potentially improve future trials in critical care.

Conclusion: The 3CT Workshop participants identified opportunities to improve critical care trials using strategies to optimize sample size calculations, account for patient and disease heterogeneity, increase the efficiency of trial conduct, maximize the use of trial data, and to refine and standardize the collection of patient-centered and patient-informed outcome measures beyond mortality.

Keywords: Critical care, Clinical trials, Acute respiratory distress syndrome, Sepsis

Randomized clinical trials (RCTs) are undertaken to identify interventions to improve clinically important outcomes using pre-specified criteria (e.g., p value < 0.05). Though several landmark critical care trials have achieved this goal, most have not [1–3]. In February 2019, a diverse group of patient representatives, regulators from the USA and European Union, federal grant managers, industry representatives, clinical trialists, epidemiologists, and clinicians convened the First *Critical Care Clinical Trialists (3CT) Workshop.* The 2-day meeting included five pre-planned sessions focused on specific critical illnesses, as well as overarching discussions focused on specific challenges in conducting and assessing critical care trials more broadly. Additional details are provided in the online supplement. The goal of the *3CT Workshop* was to share experiences, enumerate potentially modifiable trial challenges, identify shared priorities across stakeholders,

*Correspondence: mharhay@pennmedicine.upenn.edu
[1] Palliative and Advanced Illness Research (PAIR) Center, Perelman School of Medicine, University of Pennsylvania, 304 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA
Full author information is available at the end of the article
Michael O. Harhay and Jonathan D. Casey contributed equally as co-first authors.

and consolidate these into consensus recommendations and research priorities (Table 1).

The challenge

It is possible that most critical care trials have failed to demonstrate clinical benefit because the studied interventions are truly ineffective. However, there are unique challenges with critical care trials that make definitive outcome assessment difficult. For example, patients are often unable to provide consent, and surrogate decision-makers may not be available within the narrow therapeutic window of the studied intervention. Further, improved care for chronic conditions such as cancer, rheumatologic disease, cardiovascular disease, and HIV has created an increasingly complex, elderly, and immunocompromised population of critically ill patients. These pre-existing comorbidities are significant contributors to patient outcomes, but are unlikely to be modified by treatments targeting acute critical illness [4]. Finally, poorly characterized heterogeneity in patient populations and disease states may result in trials enrolling patients with opposing responses to treatment [5].

Demanding realistic trial design

Given the time and cost required to conduct a clinical trial, there is an incentive to enroll the smallest allowable number of patients. Consequently, sample size and timeline calculations have often been criticized as overly optimistic (or even unrealistic) and, therefore, predestined to produce results less informative than desired [6–9]. More worrisome, underpowered trials risk wasting resources and delivering false or uncertain conclusions about the effectiveness of tested interventions. The latter can result in the potential promotion of an ineffective therapy as well as delaying or permanently thwarting the identification of a promising therapy.

Several assumptions are required to select a target sample size for a trial. Investigators must estimate the expected control group outcome, the variation in the outcome, and the intervention-associated change. Pilot trials can be used to provide estimates of the frequency and distribution of a planned primary end point and can aid in choosing which therapies warrant further investigation by demonstrating safety, feasibility, and the therapy's effect on a surrogate in the causal pathway. However, the small size of pilot trials may cause imprecise and potentially biased estimates of treatment effects [10]. Using historical cohorts also presents challenges, as they may differ from future trial settings in patient composition and aspects of care other than the trial intervention(s). Target trial emulation, the application of design principles from randomized trials to the analysis of observational data, may provide more accurate estimates of baseline event rates and expected treatment effects [11]. However, at their core, these approaches are forms of educated guessing and may fail to reflect patient characteristics in a future trial, particularly when trials may be designed years before they enroll patients. The 3CT Workshop participants suggested building flexibility into sample size calculations when possible to allow adjustment of sample size targets mid-trial if lower than expected event rates are observed in the control group. This approach has been successfully used in several critical care trials (e.g., PROWESS-SHOCK [12] and SMART [13] trials) but requires funding agencies to allow flexible budgets and timelines so the most promising ongoing trials can be completed.

Leveraging the benefits of new trial design methodology

There is a growing interest in the application of more flexible trial designs in critical care. In this section, we discuss considerations for the trial design innovations that the 3CT Workshop participants considered most promising. A visual summary is presented in Fig. 1, and the key advantages and disadvantages of these trial design methodologies are listed in Table 2.

Prognostic enrichment

Clinical trialists have long used inclusion and exclusion criteria to create trial populations likely to benefit and unlikely to experience harm from the studied intervention, a process known as "enrichment" of the trial population. In recent years, there has been a great deal of interest in finding more sophisticated forms of trial enrichment. Prognostic enrichment designs seek to enroll patients with characteristics that suggest a high likelihood of experiencing the primary outcome. Assuming a therapy provides a consistent treatment effect to all patients, patients at the highest risk of the outcome will receive the largest absolute treatment effect. Enrolling patients with a high expected event rate facilitates greater statistical power with fewer patients. Furthermore, there may be cases where a trial intervention carries risks that might only be justified in patients with a high likelihood of a disease-related outcome [14]. Recent studies have attempted to improve upon existing risk scores and identify factors associated with disease-specific mortality that could be used for prognostic enrichment of future critical care trials [15–17]. Though attractive in principle, there are considerable limitations to this approach [18]. Patients at the highest risk of an outcome may not benefit from therapy due to the advanced stage of illness, comorbidities, or concomitant conditions. Furthermore, there are several therapies where an effect size appeared larger in the less severely ill patients (heterogeneity of treatment

Table 1 Recommendations to improve clinical trial design for critical care research from the *First Critical Care Clinical Trialists (3CT) Workshop*

Domain	Recommendation	Description and comment
Study design	Pre-specify plans for sample size re-estimation during trial design	Allows for the adjustment of the targeted sample size if outcome event rates observed in the trial differ from the initial power calculation
	Use predictive enrichment strategies for interventions in which there is a mechanistic rationale (physiologic, biologic, or genetic) to suggest why some patients may respond while others do not	Uses data from prior trials or observational data to identify patients who are likely to experience the most benefit from a given intervention, with the goal of developing enrollment criteria to selectively enroll these patients
	Use pragmatic trials to evaluate supportive therapies that might benefit a wide range of conditions or patients (e.g., early mobilization, ventilator weaning strategies, types of fluid resuscitation)	Uses broad enrollment criteria to enroll a diverse group of patients that are representative of those who would receive the intervention in usual care
	Use response-adaptive randomization for early phase trials and trials evaluating conditions with many available treatments	Incorporates information learned during the trial to (i) optimize allocation to study arms yielding the best results, which minimizes risks to patients; or (ii) optimize enrollment criteria enriching for better performing subgroups
	Evaluate opportunities to incorporate multiple trial interventions into platform trials	Simultaneously randomizes multiple, independent interventions or intervenes at multiple points in the same disease process (e.g., a trial evaluating initial therapy for a condition that feeds directly into a second trial of rescue therapies)
Study design and analysis	Incorporate a pre-specified Bayesian analysis plan with a range of priors	Analyzes trial results in the context of previously observed or presumed treatment effect distributions, producing results in terms of a likelihood of an effect on a probabilistic scale (i.e., the probability of an effect being present on a scale of 0–100)
Study conduct	Improve collaboration between critical care and pre-ICU providers (emergency medicine, pre-hospital)	Allows intervention earlier in the course of critical illness and significantly improves enrollment for interventions with narrow therapeutic windows
Outcome measures	Attempt to standardize common outcome measures across trials	Allows for meaningful across-trial comparisons
	Integrate diverse stakeholders (such as patients and families) into trial design and continue research on the development, measurement, and timing of patient-reported outcome measures	Promotes patient-centered critical care, while addressing the key challenges of patient-reported outcome measures, including the ideal timing of collection, how to account for the competing risk of mortality, and the possibility of biases introduced by incomplete long-term follow-up
Data Sharing	Encourage data sharing of de-identified patient data	Sharing data with robust data dictionaries to investigators who have pre-specified secondary analyses provides opportunities to maximize the knowledge gained from clinical trials and maximally leverages the investments made by patients, funding organizations, and researchers

effect) [19–21]. Thus, at present, there are limitations to the broad uptake of prognostic trial designs in critical care.

Predictive enrichment

Predictive designs seek to enroll patients with an increased likelihood of a benefit from a therapy, independent of the absolute risk of the outcome. Predictive enrichment is part of a broader push to provide "personalized" therapies that are chosen based on complex patient characteristics. Predictive enrichment efforts have focused on interventions in which there is a mechanistic rationale (physiologic, biologic, or genetic) suggesting why some patients may respond while others do not. Recent work has identified potential patient phenotypes of acute respiratory distress syndrome (ARDS) [22–25] and sepsis [26] which appear to experience differential responses to interventions, and several ongoing trials are using novel biomarkers to identify study populations (e.g. COMBAT-SHINE NCT04123444). A principal challenge to the use of predictive enrichment is the need for rapid, bedside identification of patient phenotypes, including the need to quantify biomarkers at the time of trial enrollment. For many interventions, it may not be practical to wait for the results of the highly specific tests needed to identify these unique phenotypes.

Pragmatic trials

Limited generalizability is an additional limitation of prognostic and predictive enrichment strategies. For supportive therapies benefitting a wide range of patients (e.g., early mobilization, ventilator weaning strategies, types of fluid resuscitation), it may be preferable to enroll a broad population of patients with few inclusion/exclusion criteria. This approach, frequently labeled as "pragmatic," attempts to enroll patients similar to those who would receive the intervention in routine care, thereby providing real-world estimation of treatment effects

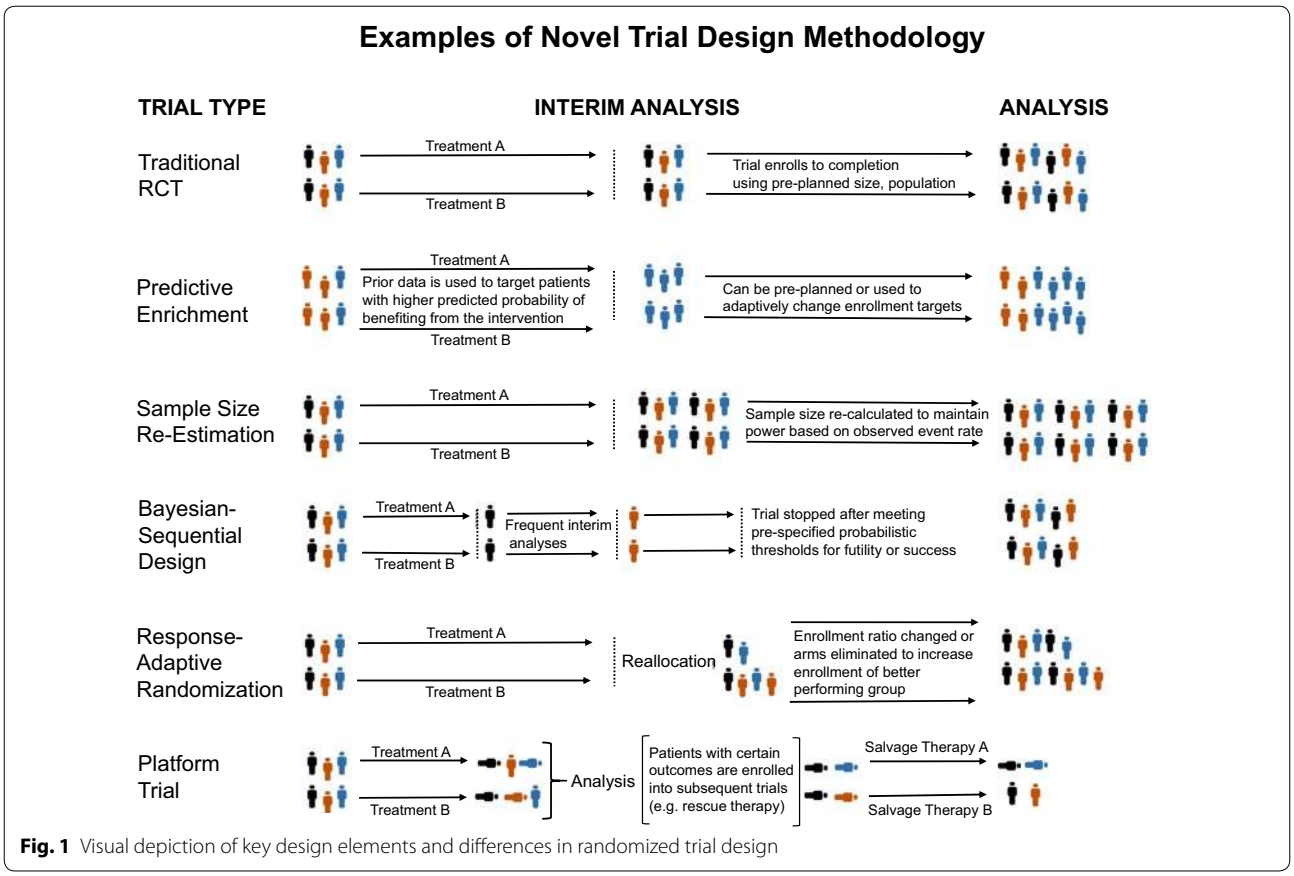

Fig. 1 Visual depiction of key design elements and differences in randomized trial design

rather than extrapolating results from highly selective explanatory trials [27]. Because heterogeneity may exist in the way patients respond to trial interventions, pragmatic trials must enroll enough patients to detect heterogeneity across diverse levels of illness severity and patient subgroups. Innovations such as the ability to easily identify patient cohorts and capture patient characteristics using electronic health records (EHRs) have facilitated the design of large pragmatic trials. Further, several recent trials have been conducted by embedding all study procedures (screening, enrollment, intervention delivery, and outcome collection) into automated tools in the EHR [13, 28, 29]. This dramatically increases study efficiency and facilitates the detection of small but clinically important treatment effects. Countries and health systems with unified EHRs have unique opportunities for multi-center, EHR-based trials [30]. Most research centers, however, are limited by the technical challenges of simultaneously accessing information from multiple, distinct EHR systems. One potential solution is the use of registry-based trials. These leverage pre-existing, prospective patient registries to identify patients and/or collect outcomes while retaining otherwise traditional trial designs [31]. Such registries could also form the foundations for platform trials (discussed below). While there is support for adopting pragmatic methodology, the delivery of complex interventions and the collection of outcomes will be a challenge for many critical care conditions.

Adaptive trials

Adaptive trial designs include plans to modify study design based on the results from trial participants (i.e., interim results). These adaptations can include changes to almost any aspect of trial design including: (i) enrollment criteria (increasing enrollment of subgroups experiencing the most benefit), (ii) sample size (preventing trials from being underpowered and stopping trials early for efficacy, safety, or futility), (iii) randomization (increasing allocation of patients to the most promising trial arm), (iv) interventions (adjusting drug doses or dropping a trial arm), and (v) analysis (converting from a non-inferiority trial to a superiority trial) (Fig. 1) [32].

Adaptive trial designs have the potential to increase efficiency, reduce costs, and increase the likelihood of identifying beneficial therapies [33]. However, there are clear challenges to adaptive trials, in addition to the increased complexity of design, conduct, and analysis.

Table 2 Advantages and disadvantages of new trial design methodologies

Design type	Advantages	Disadvantages	Examples from critical care
Prognostic enrichment	Increases trial efficiency by enrolling a population with a higher likelihood of an event, allowing adequate statistical power with fewer patients	Requires models that can reliably predict patient outcomes. Assumes a therapy provides a consistent treatment effect across the range of risks for the primary outcome; an assumption that may fail for patients at an advanced stage of illness, or for therapies that provide the largest benefit to less severely ill patients	PROWESS-SHOCK [12]
Predictive enrichment	May increase trial efficiency, reduce the impact of heterogeneity, and enhance the likelihood of identifying personalized therapies	Requires bedside methods to identify biomarkers or differentiate proposed subphenotypes prior to trial enrollment. Results may not be generalizable into the clinical setting where rapid biomarker or subphenotype identification is often difficult	AdrenOSS-2 [92] COMBAT-SHINE (NCT04123444) VIOLET [93] EUPHRATES [94]
Pragmatic trial	Maximizes generalizability and facilitates accurate effect estimates for all patients likely to receive a given intervention	Must enroll enough patients to detect heterogeneity across diverse levels of illness severity and patient subgroups	SMART [13] CRASH [95]
Sample size re-estimation (adaptive trial)	Reduces the likelihood of promising trials ending for futility and being underpowered	Requires flexible budgets and timelines. If performed blinded to treatment effect, may lead to increased expenditure of resources on ineffective treatments. If performed using treatment effects, may introduce operational biases and increase the risk of type 1 errors	PROWESS-SHOCK [12] SMART [13]
Bayesian sequential design	Increases trial efficiency by allowing early stopping for efficacy, safety, or futility. Lessens the risk of underpowered trials	Increases complexity of trial planning and execution. Difficult to predict trial duration or cost. Requires significant central effort to perform frequent analyses. May not be possible for trials with longer-term outcomes	SEPSIS-ACT [38]
Response-adaptive randomization	May increase the likelihood of identifying beneficial treatments by prospectively identifying and targeting enrollment of subgroups receiving the largest benefit or increasing treatment allocation to study arms yielding the best results, increasing power, and protecting patients	Increases the complexity of trial planning and execution. Introduces potential operational biases, as the path of trial adaptations provides insight into the outcomes of enrolled participants	PROSpect (NCT03896763)
Platform trials	Allows for more efficient conduct of clinical trials and provides the opportunity to answer multiple scientific questions with a relatively small iterative addition of effort	Dramatic increase in complexity, particularly for designs that include adaptive features. May be challenging for institutional review boards and regulatory bodies to appropriately review and oversee. Raises ethical issues regarding the ability of patients to understand full trial protocols and provide informed consent	REMAP-CAP (NCT02735707)
Bayesian trial analysis	Promotes interpretation of trial results in the context of prior research and may provide more information than dichotomous trial interpretations using a fixed *p* value	Trial interpretations may be driven by the selected prior which can be incorrect or manipulated, and for which there is not a community standard. Each prior used will result in a different trial interpretation, which can complicate decision-making and overall trial interpretations. Conducting both Bayesian and frequentist analyses increases the risk of selective reporting	EOLIA Re-analysis [67] ANDROMEDA-SHOCK Re-analysis [68]

To maintain the validity of trial results, all trial adaptations must be pre-specified with a clear design rationale and an analysis plan that controls for the potential increases to type 1 error [34]. Further, adaptive trials introduce possible operational biases as investigators may learn something about the results of enrolled patients from trial adaptions, even if the analyses are conducted in a blinded fashion. This is a particular problem in critical care research where blinding is not possible for many study interventions (physical therapy, fluid management strategy, paralysis, prone positioning, ECMO), increasing the risk of bias, where operators will observe changes in allocation ratios that modify the way they care for subsequent trial participants. To

date, adaptive trial designs have primarily been used for drug development, particularly in the field of oncology. Regulatory bodies are supportive of adaptation because of the potential benefits of minimizing the number of patients exposed to less promising treatments and shortening the time to regulatory approval [35, 36]. Many of these adaptive methodologies are applicable to clinical trials in critical care. Recent examples of adaptive design features in critical care include the RACE trial [37], an adaptive, dose-finding, phase 2 clinical trial of levocarnitine for septic shock, the SEPSIS-ACT trial of selepressin for septic shock [38], and the phase 3 PROSpect trial (NCT03896763). The PROSpect trial plans to randomize pediatric patients with ARDS to supine or prone positioning and low tidal volume or high-frequency ventilation, with adjustments in the enrollment ratios occurring every 100 patients.

Platform trials

Most clinical trials are designed to answer one specific question, but for any given patient population or disease process, there are typically many important clinical questions. Rather than committing resources to build separate trial infrastructures to screen, consent, and follow up separate patient populations for each research question, platform trials attempt to answer multiple study questions simultaneously. This can be achieved by comparing multiple interventions against a common control arm, randomizing multiple interventions for the same patient in unrelated domains, or randomizing the same patient at several different phases of their illness. Recent platform trials have begun to incorporate adaptive features ("adaptive platform trials"), allowing the addition or termination of study arms, and transition between interventions at various stages of disease. A recent example in critical care is the Randomized, Embedded, Multifactorial Adaptive Platform Trial for Community-Acquired Pneumonia (REMAP-CAP). This study randomizes patients with pneumonia to one of five antibiotic strategies to treat pneumonia, with parallel randomizations controlling the decision to give corticosteroids and the decision to give immunomodulatory macrolides (NCT02735707). While there is considerable enthusiasm regarding the potential of adaptive platform trials, these trials share the challenges of all adaptive trials and add the complexity of conducting multiple parallel interim analyses within each domain, while accounting for potential interactions among the treatment domains. Furthermore, the complex designs pose challenges for regulatory bodies and institutional review boards, and raise ethical concerns regarding whether patients can truly understand and provide informed consent for these complicated trial designs [39].

Maximizing information from clinical trials

Clinical trials demand significant investment on the part of patients, funding organizations, and researchers. There is an ethical mandate that the information gathered during clinical trials is used to provide the maximal benefit for patient care. Methods to garner additional information from clinical trials include improved pathways for data sharing, standardization of electronically abstracted data from different institutions, greater use of standardized trial definitions to support comparisons across trials, and the incorporation of pre-specified Bayesian analysis plans to augment the interpretation of trial results.

Increased data access and sharing

There has been considerable debate in recent years regarding publicly sharing de-identified data from individual clinical trial participants. The ethical and scientific arguments for data sharing are sensible and logical. Clinical trial participants have put themselves at risk, and investigators should maximize the knowledge gained from these trials. Further, clinical trial participants and funding agencies overwhelmingly favor the sharing of de-identified data to support additional research [40–42]. Therefore, in 2016, the International Committee of Medical Journal Editors (ICMJE) published a proposal requiring authors submitting clinical trial results to share de-identified individual patient data underlying trial results within 6 months of publication of the primary results [43]. Following the announcement, some researchers expressed concerns that clinical trials would be disincentivized if investigators perceived a reduced opportunity to publish important secondary analyses [44]. Others argued data sharing could produce misleading results, as subsequent researchers would be unfamiliar with the details of how data were collected [44, 45]. Finally, some noted there were significant costs and effort associated with creating de-identified data sets, and some existing data sets resulted in little additional research [46]. In response, the ICMJE retracted their proposal, instead requiring researchers to include a data sharing statement describing the opportunities for data sharing while still advocating for a long-term goal of uniform data sharing. This requirement for a data sharing plan has also been adopted by many funding bodies including the National Institutes of Health. While issues clearly remain regarding the appropriate mechanisms to protect patient privacy, acknowledge the work of clinical trialists, and protect against invalid secondary analyses, 3CT

Workshop participants broadly agreed with the importance of data sharing to advance scientific knowledge and improve patient care.

Use of standardized trial definitions

Comparing results across trials requires researchers in different settings use similar definitions. However, many common critical care outcomes lack standard definitions [47–50]. To address this problem, researchers have recently proposed 'core outcome sets' [51]. These outcome sets seek a community-wide agreement on the ideal outcome for a given disease state, its definition, and how it should be measured. To date, core outcome sets have been developed for studies evaluating interventions such as invasive mechanical ventilation [52], physical rehabilitation following critical illness [53], and extracorporeal membrane oxygenation [54]. Other core outcomes sets have focused on specific critically ill patient populations, such as the survivors of acute respiratory failure [55]. In addition to standardizing trial outcomes, maximizing the validity of across trial comparisons requires the alignment of enrollment criteria and trial interventions. A successful implementation of this approach was demonstrated by the Protocolized Resuscitation in Sepsis Meta-Analysis (PRISM) study [56], which combined data from three, recent, large clinical trials assessing early, goal-directed therapy for septic shock [57–59]. The authors prevented the biases often seen with meta-analyses by harmonizing enrollment criteria and outcome definitions and publishing a pre-specified analysis plan before pooling and unblinding trial data. Using patient-level data, they were subsequently able to demonstrate that early goal-directed therapy did not improve outcomes across a broad range of patient and care delivery factors.

Integrating Bayesian inference

Traditionally, critical care trials have been designed using a "frequentist" framework with a null and alternative hypothesis. Most trials are designed with a null hypothesis of no difference between groups, which is rejected if a difference in the outcomes between groups is sufficiently different. This design often leads to binary "positive" or "negative" trial interpretations, resulting from the dichotomous interpretation of p value thresholds to determine the existence of "statistically significant" trial results. However, a dichotomous interpretation of a p value is not inherently required (e.g., p value function [60]), and 2019 saw trends across the global scientific community to promote evaluating effect sizes and differences rather than just the use of p value thresholds. This included, for example, a compendium of articles by leading statisticians promoting new methods and views in *The American Statistician* [61], and a major statement co-signed by several hundred researchers published in *Nature* [62].

An increasingly popular alternative approach discussed by the 3CT Workshop participants, is the assessment of trial results using Bayesian inference. A key benefit of Bayesian inference is the empirical ability to formally incorporate prior knowledge when evaluating new trials. Specifically, Bayesian trial analysis involves the combining of previously observed or presumed treatment effects (called "prior" probability distributions or functions) with new trial results (called the "likelihood" function) to create a Bayesian effect estimate distribution (i.e., "posterior" probability distribution) [63, 64] (Fig. 2). Interested readers are directed to a much more comprehensive tutorial on the technical aspects and complexities of conducting a Bayesian trial analysis [64].

The differences between frequentist and Bayesian trial interpretations have been highlighted by two recent high-profile studies, the EOLIA [65] and ANDROMEDA-SHOCK trials [66]. Both trials demonstrated large, clinically significant mortality differences but failed to reach p value < 0.05. Separate post hoc Bayesian re-analyses of these trials, where the potential existence of an effect is quantified on a probabilistic scale from 0 to 100, demonstrated that the trial interventions were likely effective across a broad range of prior distributions [67, 68]. For example, in the re-analysis of the EOLIA trial even the use of a strongly skeptical prior (one which assumed an equal probability of benefit or harm) was associated with an 88% posterior probability that ECMO improved mortality [67]. Bayesian analyses cannot "fix" problems with trial design. In these two examples, highly optimistic absolute treatment-related mortality reductions > 10% were assumed along with higher than observed event rates in the control arm. However, Bayesian analysis can help ensure that trial results are more informative than a coarse "positive" or "negative" delineation, particularly in trials appearing to have clinically significant treatment differences.

There are limitations to Bayesian analyses, just as there are limitations to the popular frequentist framework. Perhaps most challenging is that each 'prior' probability is associated with a different Bayesian effect estimate (posterior) distribution, as shown in Fig. 2. Thus, a single interpretation of a trial may not be straightforward. Second, Bayesian effect estimates are heavily determined by the selected prior, which can be manipulated or incorrect, and there is not currently a standard set of agreed-upon priors for researchers to use. However, the impact of a selected prior on the posterior probability diminishes as the sample size of the newer trial increases. Further, there are strong views in the community for and against

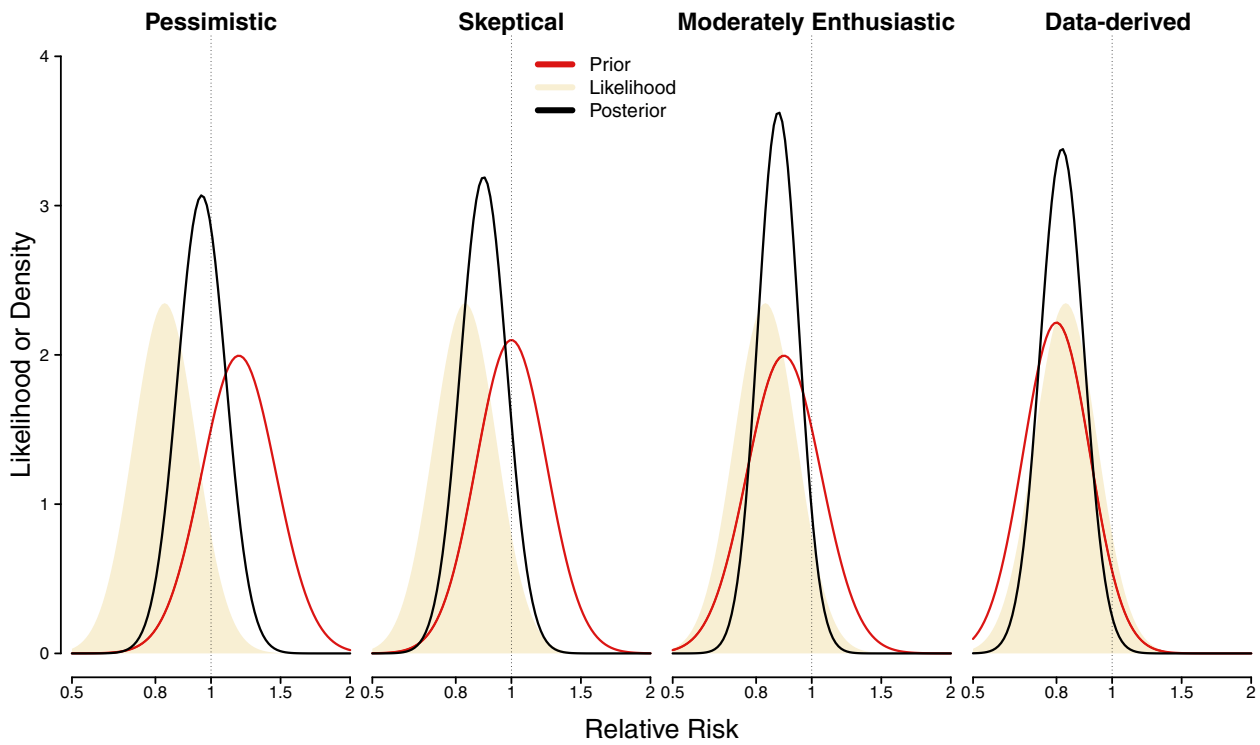

Fig. 2 Depiction of a hypothetical Bayesian trial analysis using four different prior probability distributions. To conduct a Bayesian trial analysis, researchers must first select (ideally a priori) "priors," (shown in red in the figure). Priors are meant to reflect the range of potential effect distributions that are expected before starting a trial. These priors are combined with the observed treatment effect in the trial (referred to as the likelihood function; indicated by the light brown shading). A Bayesian trial analysis typically uses a range of prior distributions, but the likelihood function is constant as it reflects the actual data observed in the trial. Each prior distribution is statistically combined with the likelihood function to create the Bayesian treatment effect estimate (i.e., posterior probability) distribution (outlined in black). As shown, the Bayesian treatment effect estimates may vary considerably based on the selected prior. There are several possible methods for choosing priors including (i) using results of observational studies, trials, or meta-analyses, (ii) eliciting expert opinion, or (iii) using a range of hypothetical distributions that assume very skeptical to enthusiastic effect distributions. The selection of priors is a critically important step in Bayesian trial analysis, and interested readers are directed to a recent technical tutorial [64], and the Bayesian re-analyses of the EOLIA [67] and ANDROMEDA-SHOCK trials [68] for guidance on prior selection and execution of a Bayesian analysis

each framework. In general, the 3CT Workshop participants agreed that Bayesian and frequentist analyses provide different but valuable information and perspectives and investigators should consider planning and reporting both frequentist and Bayesian analyses in future trials. While this promotes a less rigid, and more nuanced interpretation of trial results, this also raises the concern of selective reporting. Accordingly, the 3CT Workshop participants emphasized the importance of pre-specifying whatever analyses trialists will use to ensure the objectivity of trial interpretation.

A need for clinically meaningful, patient-centered non-mortality outcome measures

The primary outcome measure for an RCT is a critical decision as it dominates the interpretation of the trial by regulatory bodies, practicing physicians, and the broader public. The 3CT Workshop participants noted the clear relevance and importance of mortality as a safety outcome measure. However, there was a shared desire expressed, particularly from regulators and patient representatives, to incorporate patient-centered outcomes other than mortality, reflecting patients' quality of life (i.e., the challenge of surviving critical illness) in future trials.

Mortality as an outcome measure

Mortality is a logical trial outcome and many would argue survival is the most patient-centered outcome. The ideal time horizon for mortality in critical care interventions, however, remains unclear. Short-term mortality (e.g., intensive care unit [ICU], hospital, or 28-day mortality) has long been prioritized, but death is common in the period immediately following critical illness, introducing

the potential for discharge bias particularly with open-label interventions [69]. Conversely, longer-term mortality (60-day, 90-day, and 180-day) is susceptible to statistical noise from deaths unrelated to the critical illness, which may increase the number of patients needed to demonstrate a treatment effect. However, there are interventions where differences in treatment effect only become apparent with longer-term follow-up [70]. Even with short-term mortality, a significant portion can be attributed to pre-existing comorbidities and may not be modifiable [4, 71]. Further, decisions to withdraw or withhold life-sustaining treatments are common in the ICU, regardless of the effectiveness of available interventions [72–74]. One potential solution is to use disease-specific modifiable mortality in sample size estimations, but there are no reliable methods to estimate disease-specific modifiable mortality, and these calculations are likely to result in infeasibly large sample size targets [4]. These considerations have, in part, led to increased enthusiasm for non-mortality outcomes.

Non-mortality outcome measures

In recent years, there has been an increased focus on post-ICU quality-of-life, morbidity, and survivorship (e.g., post-intensive care syndrome) [75, 76]. Reductions in the number of days that patients require organ support, are in the ICU, or are in the hospital are attractive to most stakeholders and are associated with reduced long-term morbidity and mortality. A significant hurdle in greater use of these and other non-mortality outcome measures as a primary outcome is the competing risk of mortality [50, 77, 78]. When a trial participant dies, their outcomes after death are either missing (e.g., quality of life, daily organ failure, or biomarker measures) or truncated (e.g., length of stay) in a clinically relevant manner that cannot be ignored when assessing the impact of an intervention. There are effectively two general solutions available to researchers. These are (i) the use of a composite end point that combines mortality and the non-mortality outcome measure, or (ii) the analysis of the non-mortality outcome measure using an advanced statistical modeling approach to account for death [50, 78–83]. Historically, composite outcomes, such as event-free day composite measures (e.g., ventilator-free days, organ failure-free days), have been preferred by the critical care community [84, 85]. However, composite outcomes can be difficult to interpret clinically, or in other terms relevant to patients and policy makers [86–88]. There are several promising ongoing activities to develop new composite outcome measures (e.g., hierarchical end points [89]) and methods of analysis that improve upon these event-free day metrics [50, 77, 78, 90, 91]. As short-term priorities, the 3CT Workshop participants agreed that non-mortality outcomes are essential for critical care intervention assessment, but acknowledged these outcomes are more difficult to collect and analyze (particularly in the longer term). The 3CT Workshop participants also noted that establishing standardized non-mortality outcomes in a way that facilitates across-trial comparisons is a clear priority of future critical care research.

Conclusions

Significant progress has been made in the understanding of and care for patients with critical illness, but the epidemiology of critical illness is changing. Continuing advancements in critical care will require new therapies and corresponding advances in critical care trial methodology, with an emphasis on improved sample size calculations, strategic leveraging of novel trial designs, and standardization of patient-centered and patient-informed outcomes beyond mortality.

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s00134-020-05934-6) contains supplementary material, which is available to authorized users.

Author details

[1] Palliative and Advanced Illness Research (PAIR) Center, Perelman School of Medicine, University of Pennsylvania, 304 Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104-6021, USA. [2] Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [3] Division of Allergy, Pulmonary, and Critical Care Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [4] Hôpital Privé Jacques Cartier, Massy, France. [5] Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, TN, USA. [6] Department of Anesthesia, Burn and Critical Care, University Hospitals Saint-Louis - Lariboisière, AP-HP, Paris, France. [7] UMR-S 942 'MASCOT', Inserm, Paris, France. [8] Université de Paris, Paris, France. [9] Division of Critical Care Medicine, Department of Medicine, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, NY, USA. [10] Department of Anaesthesiology and Critical Care Medicine B (DAR B), Saint-Eloi Teaching Hospital, PhyMed Exp, INSERM U1046, University of Montpellier, Montpellier, France. [11] Department of Intensive Care, Cliniques St-Luc, Université Catholique de Louvain, Brussels, Belgium. [12] Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health, Toronto, Canada. [13] Cardiovascular Research Institute, University of California-San Francisco, San Francisco, CA, USA. [14] Heart Failure Society of America (HFSA), HFSA Patient Committee, Rockville, MD, USA. [15] ARDS Foundation, Northbrook, IL, USA.

Author contributions

All authors contributed to the discussions in the meeting, selection of the content for the manuscript, and the writing of the manuscript. The first draft of the manuscript was written by MOH and JDC, and all authors reviewed, edited, and commented on several previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

MOH was supported by the United States National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) grant K99 HL141678 and R00 HL141678. JDC was supported in part by the NIH/NHLBI (T32HL087738-12, K12HL133117). MAM was supported in part by NHLBI U01 HL 123004, HL140026, HL134828. MNG was supported by NHLBI UH2 HL125119, U01 HL122998, AHRQ R18 HS026188. TWR was supported in part by NHLBI U01 HL123009, NCRR U54 RR 032646, NCATS U24 TR 001608. WHS was supported in part by NHLBI U01HL123009 and NIGMS K23GM110469. JCM was supported in part by the Canadian Institutes of Health Research MYG158584.

Compliance with ethical standards

Conflicts of interest

Dr. Jaber reports receiving consulting fees from Drager, Fisher & Paykel, Medtronic, Baxter, and Fresenius-Xenios. Dr. Laterre reports personal fees from Adrenomed, Ferring, and Inotrem. Dr. Mebazaa reports personal fees from Novartis, Orion, Roche, Servier, Sanofi, Otsuka, Philips, grants and personal fees from Adrenomed, Abbott, and grants from 4TEEN4. On behalf of all authors, the corresponding author states that there are no additional potential conflicts of interest.

Publisher's Note

Received: 20 September 2019 Accepted: 11 January 2020
Published online: 18 February 2020

References

1. Santacruz CA, Pereira AJ, Celis E, Vincent JL (2019) Which multicenter randomized controlled trials in critical care medicine have shown reduced mortality? A systematic review. Crit Care Med 47(12):1680–1691. https://doi.org/10.1097/CCM.0000000000004000
2. 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. https://doi.org/10.1007/s00134-014-3272-1
3. Opal SM, Dellinger RP, Vincent JL, Masur H, Angus DC (2014) The next generation of sepsis clinical trial designs: what is next after the demise of recombinant human activated protein C? Crit Care Med 42(7):1714–1721. https://doi.org/10.1097/CCM.0000000000000325
4. Shankar-Hari M, Harrison DA, Rowan KM, Rubenfeld GD (2018) Estimating attributable fraction of mortality from sepsis to inform clinical trials. J Crit Care 45:33–39. https://doi.org/10.1016/j.jcrc.2018.01.018
5. 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. https://doi.org/10.1164/rccm.201411-2125CP
6. Harhay MO, Wagner J, Ratcliffe SJ, Bronheim RS, Gopal A, Green S, Cooney E, Mikkelsen ME, Kerlin MP, Small DS, Halpern SD (2014) Outcomes and statistical power in adult critical care randomized trials. Am J Respir Crit Care Med 189(12):1469–1478. https://doi.org/10.1164/rccm.201401-0056CP
7. Aberegg 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. https://doi.org/10.1186/cc8990
8. Rubenfeld GD (2015) Confronting the frustrations of negative clinical trials in acute respiratory distress syndrome. Ann Am Thorac Soc 12(Suppl 1):S58–63. https://doi.org/10.1513/AnnalsATS.201409-414MG
9. Anthon CT, Granholm A, Perner A, Laake JH, Moller MH (2019) Overall bias and sample sizes were unchanged in ICU trials over time: a meta-epidemiological study. J Clin Epidemiol 113:189–199. https://doi.org/10.1016/j.jclinepi.2019.05.021
10. Albers C, Lakens D (2018) When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. J Exp Soc Psychol 74:187–195
11. Admon AJ, Donnelly JP, Casey JD, Janz DR, Russell DW, Joffe AM, Vonderhaar DJ, Dischert KM, Stempek SB, Dargin JM, Rice TW, Iwashyna TJ, Semler MW (2019) Emulating a Novel Clinical Trial Using Existing Observational Data. Predicting Results of the PreVent Study. Ann Am Thorac Soc 16(8):998–1007. https://doi.org/10.1513/AnnalsATS.201903-241OC
12. Ranieri VM, Thompson BT, Barie PS, Dhainaut JF, Douglas IS, Finfer S, Gardlund B, Marshall JC, Rhodes A, Artigas A, Payen D, Tenhunen J, Al-Khalidi HR, Thompson V, Janes J, Macias WL, Vangerow B, Williams MD, Prowess-Shock Study Group (2012) Drotrecogin alfa (activated) in adults with septic shock. N Engl J Med 366(22):2055–2064. https://doi.org/10.1056/NEJMoa1202290
13. Semler MW, Self WH, Wanderer JP, Ehrenfeld JM, Wang L, Byrne DW, Stollings JL, Kumar AB, Hughes CG, Hernandez A, Guillamondegui OD, May AK, Weavind L, Casey JD, Siew ED, Shaw AD, Bernard GR, Rice TW, SMART Investigators, Pragmatic Critical Care Research Group (2018) Balanced crystalloids versus saline in critically Ill adults. N Engl J Med 378(9):829–839. https://doi.org/10.1056/NEJMoa1711584
14. Wong HR (2017) Intensive care medicine in 2050: precision medicine. Intensive Care Med 43(10):1507–1509. https://doi.org/10.1007/s00134-017-4727-y
15. Villar J, Ambros A, Mosteiro F, Martinez D, Fernandez L, Ferrando C, Carriedo D, Soler JA, Parrilla D, Hernandez M, Andaluz-Ojeda D, Anon JM, Vidal A, Gonzalez-Higueras E, Martin-Rodriguez C, Diaz-Lamas AM, Blanco J, Belda J, Diaz-Dominguez FJ, Rico-Feijoo J, Martin-Delgado C, Romera MA, Gonzalez-Martin JM, Fernandez RL, Kacmarek RM, Spanish Initiative for Epidemiology, Stratification, and Therapies of ARDS (SIESTA) Network (2019) A prognostic enrichment strategy for selection of patients with acute respiratory distress syndrome in clinical trials. Crit Care Med 47(3):377–385. https://doi.org/10.1097/CCM.0000000000003624
16. Langley RJ, Tsalik EL, van Velkinburgh JC, Glickman SW, Rice BJ, Wang C, Chen B, Carin L, Suarez A, Mohney RP, Freeman DH, Wang M, You J, Wulff J, Thompson JW, Moseley MA, Reisinger S, Edmonds BT, Grinnell B, Nelson DR, Dinwiddie DL, Miller NA, Saunders CJ, Soden SS, Rogers AJ, Gazourian L, Fredenburgh LE, Massaro AF, Baron RM, Choi AM, Corey GR, Ginsburg GS, Cairns CB, Otero RM, Fowler VG Jr, Rivers EP, Woods CW, Kingsmore SF (2013) An integrated clinico-metabolomic model improves prediction of death in sepsis. Sci Transl Med 5(195):195ra95. https://doi.org/10.1126/scitranslmed.3005893
17. Wong HR, Atkinson SJ, Cvijanovich NZ, Anas N, Allen GL, Thomas NJ, Bigham MT, Weiss SL, Fitzgerald JC, Checchia PA, Meyer K, Quasney M, Hall M, Gedeit R, Freishtat RJ, Nowak J, Raj SS, Gertz S, Lindsell CJ (2016) Combining prognostic and predictive enrichment strategies to identify children with septic shock responsive to corticosteroids. Crit Care Med 44(10):e1000–1003. https://doi.org/10.1097/CCM.0000000000001833
18. Vincent JL, Opal SM, Marshall JC (2010) Ten reasons why we should NOT use severity scores as entry criteria for clinical trials or in our treatment decisions. Crit Care Med 38(1):283–287. https://doi.org/10.1097/CCM.0b013e3181b785a2
19. Russell JA, Walley KR, Singer J, Gordon AC, Hebert PC, Cooper DJ, Holmes CL, Mehta S, Granton JT, Storms MM, Cook DJ, Presneill JJ, Ayers D, VASST Investigators (2008) Vasopressin versus norepinephrine infusion in patients with septic shock. N Engl J Med 358(9):877–887. https://doi.org/10.1056/NEJMoa067373
20. Krag M, Marker S, Perner A, Wetterslev J, Wise MP, Schefold JC, Keus F, Guttormsen AB, Bendel S, Borthwick M, Lange T, Rasmussen BS, Siegemund M, Bundgaard H, Elkmann T, Jensen JV, Nielsen RD, Liboriussen L, Bestle MH, Elkjaer JM, Palmqvist DF, Backlund M, Laake JH, Badstolokken PM, Gronlund J, Breum O, Walli A, Winding R, Iversen S, Jarnvig IL, White JO, Brand B, Madsen MB, Quist L, Thornberg KJ, Moller A, Wiis J, Granholm A, Anthon CT, Meyhoff TS, Hjortrup PB, Aagaard SR, Andreasen JB, Sorensen CA, Haure P, Hauge J, Hollinger A, Scheuzger J, Tuchscherer D, Vuilliomenet T, Takala J, Jakob SM, Vang ML, Paelestik KB, Andersen KLD, van der Horst ICC, Dieperink W, Fjolner J, Kjer CKW, Solling C, Solling CG, Karttunen J, Morgan MPG, Sjobo B, Engstrom J, Agerholm-Larsen B, Moller MH, SUP-ICU Trial Group (2018) Pantoprazole in patients at risk for gastrointestinal bleeding in the ICU. N Engl J Med 379(23):2199–2208. https://doi.org/10.1056/NEJMoa1714919
21. Marker S, Perner A, Wetterslev J, Krag M, Lange T, Wise MP, Borthwick M, Bendel S, Keus F, Guttormsen AB, Schefold JC, Moller MH, SUP-ICU Investigators (2019) Pantoprazole prophylaxis in ICU patients with high severity of disease: a post hoc analysis of the placebo-controlled SUP-ICU trial. Intensive Care Med 45(5):609–618. https://doi.org/10.1007/s00134-019-05589-y
22. Calfee CS, Delucchi K, Parsons PE, Thompson BT, Ware LB, Matthay MA, NHLBI ARDS Network (2014) Subphenotypes in acute respiratory distress syndrome: latent class analysis of data from two randomised controlled trials. Lancet Respir Med 2(8):611–620. https://doi.org/10.1016/S2213-2600(14)70097-9
23. Delucchi K, Famous KR, Ware LB, Parsons PE, Thompson BT, Calfee CS, NHLBI ARDS Network (2018) Stability of ARDS subphenotypes over time in two randomised controlled trials. Thorax 73(5):439–445. https://doi.org/10.1136/thoraxjnl-2017-211090

24. Famous KR, Delucchi K, Ware LB, Kangelaris KN, Liu KD, Thompson BT, Calfee CS, NHLBI ARDS Network (2017) Acute respiratory distress syndrome subphenotypes respond differently to randomized fluid management strategy. Am J Respir Crit Care Med 195(3):331–338. https://doi.org/10.1164/rccm.201603-0645OC
25. Calfee CS, Delucchi KL, Sinha P, Matthay MA, Hackett J, Shankar-Hari M, McDowell C, Laffey JG, O'Kane CM, McAuley DF, Irish Critical Care Trials Group (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. https://doi.org/10.1016/S2213-2600(18)30177-2
26. Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, Berry S, Clermont G, Cooper G, Gomez H, Huang DT, Kellum JA, Mi Q, Opal SM, Talisa V, van der Poll T, Visweswaran S, Vodovotz Y, Weiss JC, Yealy DM, Yende S, Angus DC (2019) Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA 321(20):2003–2017. https://doi.org/10.1001/jama.2019.5791
27. Bellomo R, Landoni G, Young P (2016) Improved survival in critically ill patients: are large RCTs more useful than personalized medicine? Yes. Intensive Care Med 42(11):1775–1777. https://doi.org/10.1007/s00134-016-4491-4
28. Semler MW, Wanderer JP, Ehrenfeld JM, Stollings JL, Self WH, Siew ED, Wang L, Byrne DW, Shaw AD, Bernard GR, Rice TW, SALT Investigators, Pragmatic Critical Care Research Group (2017) Balanced Crystalloids versus Saline in the Intensive Care Unit The SALT Randomized Trial. Am J Respir Crit Care Med 195(10):1362–1372. https://doi.org/10.1164/rccm.201607-1345OC
29. Self WH, Semler MW, Wanderer JP, Wang L, Byrne DW, Collins SP, Slovis CM, Lindsell CJ, Ehrenfeld JM, Siew ED, Shaw AD, Bernard GR, Rice TW, SALT-ED Investigators (2018) Balanced crystalloids versus saline in noncritically Ill adults. N Engl J Med 378(9):819–828. https://doi.org/10.1056/NEJMoa1711586
30. Shah ASV, Anand A, Strachan FE, Ferry AV, Lee KK, Chapman AR, Sandeman D, Stables CL, Adamson PD, Andrews JPM, Anwar MS, Hung J, Moss AJ, O'Brien R, Berry C, Findlay I, Walker S, Cruickshank A, Reid A, Gray A, Collinson PO, Apple FS, McAllister DA, Maguire D, Fox KAA, Newby DE, Tuck C, Harkess R, Parker RA, Keerie C, Weir CJ, Mills NL, High-STEACS Investigators (2018) High-sensitivity troponin in the evaluation of patients with suspected acute coronary syndrome: a stepped-wedge, cluster-randomised controlled trial. Lancet 392(10151):919–928. https://doi.org/10.1016/S0140-6736(18)31923-8
31. Frobert O, Lagerqvist B, Olivecrona GK, Omerovic E, Gudnason T, Maeng M, Aasa M, Angeras O, Calais F, Danielewicz M, Erlinge D, Hellsten L, Jensen U, Johansson AC, Karegren A, Nilsson J, Robertson L, Sandhall L, Sjogren I, Ostlund O, Harnek J, James SK (2013) Thrombus aspiration during ST-segment elevation myocardial infarction. N Engl J Med 369(17):1587–1597. https://doi.org/10.1056/NEJMoa1308789
32. van Werkhoven CH, Harbarth S, Bonten MJM (2019) Adaptive designs in clinical trials in critically ill patients: principles, advantages and pitfalls. Intensive Care Med 45(5):678–682. https://doi.org/10.1007/s00134-018-5426-z
33. Bhatt DL, Mehta C (2016) Adaptive designs for clinical trials. N Engl J Med 375(1):65–74. https://doi.org/10.1056/NEJMra1510061
34. Posch M, Maurer W, Bretz F (2011) Type I error rate control in adaptive designs for confirmatory clinical trials with treatment selection at interim. Pharm Stat 10(2):96–104. https://doi.org/10.1002/pst.413
35. Elsasser A, Regnstrom J, Vetter T, Koenig F, Hemmings RJ, Greco M, Papaluca-Amati M, Posch M (2014) Adaptive clinical trial designs for European marketing authorization: a survey of scientific advice letters from the European Medicines Agency. Trials 15:383. https://doi.org/10.1186/1745-6215-15-383
36. US Food and Drug Administration (FDA). Guidance: Adaptive Design Clinical Trials for Drugs and Biologics [Draft Guidance current as of October 31, 2018]. Docket Number FDA-2018-D-3124. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/adaptive-design-clinical-trials-drugs-and-biologics. Accessed Nov 22 2019
37. Jones AE, Puskarich MA, Shapiro NI, Guirgis FW, Runyon M, Adams JY, Sherwin R, Arnold R, Roberts BW, Kurz MC, Wang HE, Kline JA, Courtney DM, Trzeciak S, Sterling SA, Nandi U, Patki D, Viele K (2018) Effect of Levocarnitine vs placebo as an adjunctive treatment for septic shock: the Rapid Administration of Carnitine in Sepsis (RACE) randomized clinical trial. JAMA Netw Open 1(8):e186076. https://doi.org/10.1001/jamanetworkopen.2018.6076
38. Laterre PF, Berry SM, Blemings A, Carlsen JE, Francois B, Graves T, Jacobsen K, Lewis RJ, Opal SM, Perner A, Pickkers P, Russell JA, Windelov NA, Yealy DM, Asfar P, Bestle MH, Muller G, Bruel C, Brule N, Decruyenaere J, Dive AM, Dugernier T, Krell K, Lefrant JY, Megarbane B, Mercier E, Mira JP, Quenot JP, Rasmussen BS, Thorsen-Meyer HC, Vander Laenen M, Vang ML, Vignon P, Vinatier I, Wichmann S, Wittebole X, Kjolbye AL, Angus DC, SEPSIS-ACT Investigators (2019) Effect of selepressin vs placebo on ventilator- and vasopressor-free days in patients with septic shock: The SEPSIS-ACT randomized clinical trial. JAMA 322(15):1476–1485. https://doi.org/10.1001/jama.2019.14607
39. Adaptive Platform Trials Coalition (2019) Adaptive platform trials: definition, design, conduct and reporting considerations. Nat Rev Drug Discov 18(10):797–807. https://doi.org/10.1038/s41573-019-0034-3
40. Mello MM, Lieou V, Goodman SN (2018) Clinical trial participants' views of the risks and benefits of data sharing. N Engl J Med 378(23):2202–2211. https://doi.org/10.1056/NEJMsa1713258
41. Haug CJ (2017) Whose data are they anyway? Can a patient perspective advance the data-sharing debate? N Engl J Med 376(23):2203–2205. https://doi.org/10.1056/NEJMp1704485
42. Kiley R, Peatfield T, Hansen J, Reddington F (2017) Data sharing from clinical trials–a research funder's perspective. N Engl J Med 377(20):1990–1992. https://doi.org/10.1056/NEJMsb1708278
43. Taichman DB, Backus J, Baethge C, Bauchner H, de Leeuw PW, Drazen JM, Fletcher J, Frizelle FA, Groves T, Haileamlak A, James A, Laine C, Peiperl L, Pinborg A, Sahni P, Wu S (2016) Sharing clinical trial data: a proposal from the international committee of medical journal editors. PLoS Med 13(1):e1001950. https://doi.org/10.1371/journal.pmed.1001950
44. Rosenbaum L (2017) Bridging the data-sharing divide - seeing the devil in the details, not the other Camp. N Engl J Med 376(23):2201–2203. https://doi.org/10.1056/NEJMp1704482
45. International Consortium of Investigators for Fairness in Trial Data Sharing, Devereaux PJ, Guyatt G, Gerstein H, Connolly S, Yusuf S (2016) Toward Fairness in Data Sharing. N Engl J Med 375(5):405–407. https://doi.org/10.1056/NEJMp1605654
46. Strom BL, Buyse ME, Hughes J, Knoppers BM (2016) Data sharing–Is the juice worth the squeeze? N Engl J Med 375(17):1608–1609. https://doi.org/10.1056/NEJMp1610336
47. Contentin L, Ehrmann S, Giraudeau B (2014) Heterogeneity in the definition of mechanical ventilation duration and ventilator-free days. Am J Respir Crit Care Med 189(8):998–1002. https://doi.org/10.1164/rccm.201308-1499LE
48. Blackwood B, Clarke M, McAuley DF, McGuigan PJ, Marshall JC, Rose L (2014) How outcomes are defined in clinical trials of mechanically ventilated adults and children. Am J Respir Crit Care Med 189(8):886–893. https://doi.org/10.1164/rccm.201309-1645PP
49. Bourcier S, Hindlet P, Guidet B, Dechartres A (2019) Reporting of organ support outcomes in septic shock randomized controlled trials: a methodologic review-the sepsis organ support study. Crit Care Med 47(7):984–992. https://doi.org/10.1097/CCM.0000000000003746
50. Harhay MO, Ratcliffe SJ, Small DS, Suttner LH, Crowther MJ, Halpern SD (2019) Measuring and analyzing length of stay in critical care trials. Med Care 57(9):e53–e59. https://doi.org/10.1097/MLR.0000000000001059
51. Blackwood B, Marshall J, Rose L (2015) Progress on core outcome sets for critical care research. Curr Opin Crit Care 21(5):439–444. https://doi.org/10.1097/MCC.0000000000000232
52. Blackwood B, Ringrow S, Clarke M, Marshall JC, Connolly B, Rose L, McAuley DF (2019) A core outcome set for critical care ventilation trials. Crit Care Med 47(10):1324–1331. https://doi.org/10.1097/CCM.0000000000003904
53. Connolly B, Denehy L, Hart N, Pattison N, Williamson P, Blackwood B (2018) Physical rehabilitation core outcomes in critical illness (PRACTICE): protocol for development of a core outcome set. Trials 19(1):294. https://doi.org/10.1186/s13063-018-2678-4
54. Hodgson CL, Burrell AJC, Engeler DM, Pellegrino VA, Brodie D, Fan E, International ECMO Network (2019) Core Outcome Measures for Research in Critically Ill Patients Receiving Extracorporeal Membrane Oxygenation for Acute Respiratory or Cardiac Failure: An International, Multidisciplinary, Modified Delphi Consensus Study. Crit Care Med 47(11):1557–1563. https://doi.org/10.1097/CCM.0000000000003954

55. Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO 3rd, Turnbull AE (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. https://doi.org/10.1164/rccm.201702-0372OC
56. PRISM Investigators, Rowan KM, Angus DC, Bailey M, Barnato AE, Bellomo R, Canter RR, Coats TJ, Delaney A, Gimbel E, Grieve RD, Harrison DA, Higgins AM, Howe B, Huang DT, Kellum JA, Mouncey PR, Music E, Peake SL, Pike F, Reade MC, Sadique MZ, Singer M, Yealy DM (2017) Early, goal-directed therapy for septic shock—a patient-level meta-analysis. N Engl J Med 376(23):2223–2234. https://doi.org/10.1056/NEJMoa1701380
57. ProCess Investigators, Yealy DM, Kellum JA, Huang DT, Barnato AE, Weissfeld LA, Pike F, Terndrup T, Wang HE, Hou PC, LoVecchio F, Filbin MR, Shapiro NI, Angus DC (2014) A randomized trial of protocol-based care for early septic shock. N Engl J Med 370(18):1683–1693. https://doi.org/10.1056/NEJMoa1401602
58. ARISE Investigators, ANZICS Clinical Trials Group, Peake SL, Delaney A, Bailey M, Bellomo R, Cameron PA, Cooper DJ, Higgins AM, Holdgate A, Howe BD, Webb SA, Williams P (2014) Goal-directed resuscitation for patients with early septic shock. N Engl J Med 371(16):1496–1506. https://doi.org/10.1056/NEJMoa1404380
59. Mouncey PR, Osborn TM, Power GS, Harrison DA, Sadique MZ, Grieve RD, Jahan R, Harvey SE, Bell D, Bion JF, Coats TJ, Singer M, Young JD, Rowan KM, ProMISe Trial Investigators (2015) Trial of early, goal-directed resuscitation for septic shock. N Engl J Med 372(14):1301–1311. https://doi.org/10.1056/NEJMoa1500896
60. Infanger D, Schmidt-Trucksass A (2019) P value functions: An underused method to present research results and to promote quantitative reasoning. Stat Med 38(21):4189–4197. https://doi.org/10.1002/sim.8293
61. Wasserstein RL, Schirm AL, Lazar NA (2019) Moving to a World Beyond "p, 0.05". Am Stat 73(1):1–19. https://doi.org/10.1080/00031305.2019.1583913
62. Amrhein V, Greenland S, McShane B (2019) Scientists rise up against statistical significance. Nature 567(7748):305–307. https://doi.org/10.1038/d41586-019-00857-9
63. Lewis RJ, Angus DC (2018) Time for clinicians to embrace their inner bayesian?: reanalysis of results of a clinical trial of extracorporeal membrane oxygenation. JAMA 320(21):2208–2210. https://doi.org/10.1001/jama.2018.16916
64. Bittl JA, He Y (2017) Bayesian analysis: A practical approach to interpret clinical trials and create clinical practice guidelines. Circ Cardiovasc Qual Outcomes 10(8):9–10. https://doi.org/10.1161/CIRCOUTCOMES.117.003563
65. Combes A, Hajage D, Capellier G, Demoule A, Lavoue S, Guervilly C, Da Silva D, Zafrani L, Tirot P, Veber B, Maury E, Levy B, Cohen Y, Richard C, Kalfon P, Bouadma L, Mehdaoui H, Beduneau G, Lebreton G, Brochard L, Ferguson ND, Fan E, Slutsky AS, Brodie D, Mercat A, EOLIA Trial Group, REVA, ECMONet (2018) Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. N Engl J Med 378(21):1965–1975. https://doi.org/10.1056/NEJMoa1800385
66. Hernandez G, Ospina-Tascon GA, Damiani LP, Estenssoro E, Dubin A, Hurtado J, Friedman G, Castro R, Alegria L, Teboul JL, Cecconi M, Ferri G, Jibaja M, Pairumani R, Fernandez P, Barahona D, Granda-Luna V, Cavalcanti AB, Bakker J, Hernandez G, Ospina-Tascon G, Petri Damiani L, Estenssoro E, Dubin A, Hurtado J, Friedman G, Castro R, Alegria L, Teboul JL, Cecconi M, Cecconi M, Ferri G, Jibaja M, Pairumani R, Fernandez P, Barahona D, Cavalcanti AB, Bakker J, ANDROMEDA SHOCK Investigators, and the Latin America Intensive Care Network (LIVEN), Hernandez G, Alegria L, Ferri G, Rodriguez N, Holger P, Soto N, Pozo M, Bakker J, Cook D, Vincent JL, Rhodes A, Kavanagh BP, Dellinger P, Rietdijk W, Carpio D, Pavez N, Henriquez E, Bravo S, Valenzuela ED, Vera M, Dreyse J, Oviedo V, Cid MA, Larroulet M, Petruska E, Sarabia C, Gallardo D, Sanchez JE, Gonzalez H, Arancibia JM, Munoz A, Ramirez G, Aravena F, Aquevedo A, Zambrano F, Bozinovic M, Valle F, Ramirez M, Rossel V, Munoz P, Ceballos C, Esveile C, Carmona C, Candia E, Mendoza D, Sanchez A, Ponce D, Ponce D, Lastra J, Nahuelpan B, Fasce F, Luengo C, Medel N, Cortes C, Campassi L, Rubatto P, Horna N, Furche M, Pendino JC, Bettini L, Lovesio C, Gonzalez MC, Rodruguez J, Canales H, Caminos F, Galletti C, Minoldo E, Aramburu MJ, Olmos D, Nin N, Tenzi J, Quiroga C, Lacuesta P, Gaudin A, Pais R, Silvestre A, Olivera G, Rieppi G, Berrutti D, Ochoa M, Cobos P, Vintimilla F, Ramirez V, Tobar M, Garcia F, Picoita F, Remache N, Granda V, Paredes F, Barzallo E, Garces P, Guerrero F, Salazar S, Torres G, Tana C, Calahorrano J, Solis F, Torres P, Herrera L, Ornes A, Perez V, Delgado G, Lopez A, Espinosa E, Moreira J, Salcedo B, Villacres I, Suing J, Lopez M, Gomez L, Toctaquiza G, Cadena Zapata M, Orazabal MA, Pardo Espejo R, Jimenez J, Calderon A, Paredes G, Barberan JL, Moya T, Atehortua H, Sabogal R, Ortiz G, Lara A, Sanchez F, Hernan Portilla A, Davila H, Mora JA, Calderon LE, Alvarez I, Escobar E, Bejarano A, Bustamante LA, Aldana JL (2019) Effect of a Resuscitation strategy targeting peripheral perfusion status vs serum lactate levels on 28-day mortality among patients with septic shock: the ANDROMEDA-SHOCK randomized clinical trial. JAMA 321(7):654–664. https://doi.org/10.1001/jama.2019.0071
67. Goligher EC, Tomlinson G, Hajage D, Wijeysundera DN, Fan E, Juni P, Brodie D, Slutsky AS, Combes A (2018) Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome and posterior probability of mortality benefit in a post Hoc bayesian analysis of a randomized clinical trial. JAMA 320(21):2251–2259. https://doi.org/10.1001/jama.2018.14276
68. Zampieri FG, Damiani LP, Bakker J, Ospina-Tascon GA, Castro R, Cavalcanti AB, Hernandez G, ANDROMEDA-SHOCK Investigators, Latin America Intensive Care Network (LIVEN) (2019) Effects of a resuscitation strategy targeting peripheral perfusion status vs serum lactate levels among patients with septic shock: a bayesian reanalysis of the ANDROMEDA-SHOCK trial. Am J Respir Crit Care Med. https://doi.org/10.1164/rccm.201905-0968OC
69. Vasilevskis EE, Kuzniewicz MW, Dean ML, Clay T, Vittinghoff E, Rennie DJ, Dudley RA (2009) Relationship between discharge practices and intensive care unit in-hospital mortality performance: evidence of a discharge bias. Med Care 47(7):803–812. https://doi.org/10.1097/MLR.0b013e3181a39454
70. Perner A, Haase N, Guttormsen AB, Tenhunen J, Klemenzson G, Aneman A, Madsen KR, Moller MH, Elkjaer JM, Poulsen LM, Bendtsen A, Winding R, Steensen M, Berezowicz P, Soe-Jensen P, Bestle M, Strand K, Wiis J, White JO, Thornberg KJ, Quist L, Nielsen J, Andersen LH, Holst LB, Thormar K, Kjaeldgaard AL, Fabritius ML, Mondrup F, Pott FC, Moller TP, Winkel P, Wetterslev J, 6S Trial Group, Scandinavian Critical Care Trials Group (2012) Hydroxyethyl starch 130/0.42 versus Ringer's acetate in severe sepsis. N Engl J Med 367(2):124–134
71. Rhee C, Jones TM, Hamad Y, Pande A, Varon J, O'Brien C, Anderson DJ, Warren DK, Dantes RB, Epstein L, Klompas M, Centers for Disease Control and Prevention (CDC) Prevention Epicenters Program (2019) Prevalence, underlying causes, and preventability of sepsis-associated mortality in US acute care hospitals. JAMA Netw Open 2(2):e187571. https://doi.org/10.1001/jamanetworkopen.2018.7571
72. Messika J, Gaudry S, Tubach F, Guillo S, Dreyfuss D, Hajage D, Ricard JD (2018) Underreporting of end-of-life decisions in critical care trials: a call to modify the consolidated standards of reporting trials statement. Am J Respir Crit Care Med 197(2):263–266. https://doi.org/10.1164/rccm.201703-0586LE
73. Mehter HM, Wiener RS, Walkey AJ (2014) "Do not resuscitate" decisions in acute respiratory distress syndrome. A secondary analysis of clinical trial data. Ann Am Thorac Soc 11(10):1592–1596. https://doi.org/10.1513/AnnalsATS.201406-244BC
74. Lautrette A, Garrouste-Orgeas M, Bertrand PM, Goldgran-Toledano D, Jamali S, Laurent V, Argaud L, Schwebel C, Mourvillier B, Darmon M, Ruckly S, Dumenil AS, Lemiale V, Souweine B, Timsit JF, Outcomerea Study Group (2015) Respective impact of no escalation of treatment, withholding and withdrawal of life-sustaining treatment on ICU patients' prognosis: a multicenter study of the Outcomerea Research Group. Intensive Care Med 41(10):1763–1772. https://doi.org/10.1007/s00134-015-3944-5
75. Bein T, Bienvenu OJ, Hopkins RO (2019) Focus on long-term cognitive, psychological and physical impairments after critical illness. Intensive Care Med 45(10):1466–1468. https://doi.org/10.1007/s00134-019-05718-7
76. Dinglas VD, Faraone LN, Needham DM (2018) Understanding patient-important outcomes after critical illness: a synthesis of recent qualitative, empirical, and consensus-related studies. Curr Opin Crit Care 24(5):401–409. https://doi.org/10.1097/MCC.0000000000000533
77. Timsit JF, de Kraker MEA, Sommer H, Weiss E, Bettiol E, Wolkewitz M, Nikolakopoulos S, Wilson D, Harbarth S, COMBACTE-NET consortium

(2017) Appropriate endpoints for evaluation of new antibiotic therapies for severe infections: a perspective from COMBACTE's STAT-Net. Intensive Care Med 43(7):1002–1012. https://doi.org/10.1007/s00134-017-4802-4
78. Colantuoni E, Scharfstein DO, Wang C, Hashem MD, Leroux A, Needham DM, Girard TD (2018) Statistical methods to compare functional outcomes in randomized controlled trials with high mortality. BMJ 360:j5748. https://doi.org/10.1136/bmj.j5748
79. Colantuoni E, Dinglas VD, Ely EW, Hopkins RO, Needham DM (2016) Statistical methods for evaluating delirium in the ICU. Lancet Respir Med 4(7):534–536. https://doi.org/10.1016/S2213-2600(16)30138-2
80. Hayden D, Pauler DK, Schoenfeld D (2005) An estimator for treatment comparisons among survivors in randomized trials. Biometrics 61(1):305–310. https://doi.org/10.1111/j.0006-341X.2005.030227.x
81. Chiba Y, VanderWeele TJ (2011) A simple method for principal strata effects when the outcome has been truncated due to death. Am J Epidemiol 173(7):745–751. https://doi.org/10.1093/aje/kwq418
82. Deslandes E, Chevret S (2010) Joint modeling of multivariate longitudinal data and the dropout process in a competing risk setting: application to ICU data. BMC Med Res Methodol 10:69. https://doi.org/10.1186/1471-2288-10-69
83. Checkley W, Brower RG, Munoz A, NHLBI ARDS Network (2010) Inference for mutually exclusive competing events through a mixture of generalized gamma distributions. Epidemiology 21(4):557–565. https://doi.org/10.1097/EDE.0b013e3181e090ed
84. Schoenfeld DA, Bernard GR, NHLBI ARDS Network (2002) Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. Crit Care Med 30(8):1772–1777. https://doi.org/10.1097/00003246-200208000-00016
85. Young P, Hodgson C, Dulhunty J, Saxena M, Bailey M, Bellomo R, Davies A, Finfer S, Kruger P, Lipman J, Myburgh J, Peake S, Seppelt I, Streat S, Tate R, Webb S, ANZICS Clinical Trials Group (2012) End points for phase II trials in intensive care: recommendations from the Australian and New Zealand Clinical Trials Group consensus panel meeting. Crit Care Resusc 14(3):211–215
86. Tomlinson G, Detsky AS (2010) Composite end points in randomized trials: there is no free lunch. JAMA 303(3):267–268. https://doi.org/10.1001/jama.2009.2017
87. Bodet-Contentin L, Frasca D, Tavernier E, Feuillet F, Foucher Y, Giraudeau B (2018) Ventilator-free day outcomes can be misleading. Crit Care Med 46(3):425–429. https://doi.org/10.1097/CCM.0000000000002890
88. Yehya N, Harhay MO, Curley MAQ, Schoenfeld DA, Reeder RW (2019) Reappraisal of ventilator-free days in critical care research. Am J Respir Crit Care Med 200(7):828–836. https://doi.org/10.1164/rccm.201810-2050CP
89. Novack V, Beitler JR, Yitshak-Sade M, Thompson BT, Schoenfeld DA, Rubenfeld G, Talmor D, Brown SM (2020) Alive and ventilator free: a hierarchical, composite outcome for clinical trials in the acute respiratory distress syndrome. Crit Care Med 48(2):158–166. https://doi.org/10.1097/CCM.0000000000004104
90. Wang C, Scharfstein DO, Colantuoni E, Girard TD, Yan Y (2017) Inference in randomized trials with death and missingness. Biometrics 73(2):431–440. https://doi.org/10.1111/biom.12594
91. Lin W, Halpern SD, Prasad Kerlin M, Small DS (2017) A "placement of death" approach for studies of treatment effects on ICU length of stay. Stat Methods Med Res 26(1):292–311. https://doi.org/10.1177/0962280214545121
92. Geven C, Blet A, Kox M, Hartmann O, Scigalla P, Zimmermann J, Marx G, Laterre PF, Mebazaa A, Pickkers P (2019) A double-blind, placebo-controlled, randomised, multicentre, proof-of-concept and dose-finding phase II clinical trial to investigate the safety, tolerability and efficacy of adrecizumab in patients with septic shock and elevated adrenomedullin concentration (AdrenOSS-2). BMJ Open 9(2):e024475. https://doi.org/10.1136/bmjopen-2018-024475
93. NHLBI PETAL Clinical Trials Network, Ginde AA, Brower RG, Caterino JM, Finck L, Banner-Goodspeed VM, Grissom CK, Hayden D, Hough CL, Hyzy RC, Khan A, Levitt JE, Park PK, Ringwood N, Rivers EP, Self WH, Shapiro NI, Thompson BT, Yealy DM, Talmor D (2019) Early high-dose vitamin D3 for critically Ill, vitamin D-deficient patients. N Engl J Med 381(26):2529–2540. https://doi.org/10.1056/NEJMoa1911124
94. Dellinger RP, Bagshaw SM, Antonelli M, Foster DM, Klein DJ, Marshall JC, Palevsky PM, Weisberg LS, Schorr CA, Trzeciak S, Walker PM, EUPHRATES Trial Investigators (2018) Effect of targeted polymyxin B hemoperfusion on 28-day mortality in patients with septic shock and elevated endotoxin level: the EUPHRATES randomized clinical trial. JAMA 320(14):1455–1463. https://doi.org/10.1001/jama.2018.14618
95. Roberts I, Yates D, Sandercock P, Farrell B, Wasserberg J, Lomas G, Cottingham R, Svoboda P, Brayley N, Mazairac G, Laloe V, Munoz-Sanchez A, Arango M, Hartzenberg B, Khamis H, Yutthakasemsunt S, Komolafe E, Olldashi F, Yadav Y, Murillo-Cabezas F, Shakur H, Edwards P, CRASH Trial Collaborators (2004) Effect of intravenous corticosteroids on death within 14 days in 10008 adults with clinically significant head injury (MRC CRASH trial): randomised placebo-controlled trial. Lancet 364(9442):1321–1328. https://doi.org/10.1016/S0140-6736(04)17188-2