Important issues in the design and reporting of clinical trials in severe sepsis and acute lung injury

R. Phillip Dellinger MDa,⁎, Jean-Louis Vincent MD, PhDb, John Marshall MDc, Konrad Reinhart MDD

aCooper University Hospital, Robert Wood Johnson Medical School, Camden, NJ, USA
bDepartment of Intensive Care, Erasme University Hospital, Université Libre de Bruxelles, Belgium
cDepartments of Surgery and Critical Care Medicine, University of Toronto, Li Ka Shing Knowledge Institute, St Michael’s Hospital, Toronto, Canada
dUniversity Hospital Jena, Friedrich Schiller University, Jena, Germany

Keywords:
Clinical trials;
Severe sepsis;
Acute lung injury

Abstract
Severe sepsis and acute lung injury are challenging diagnoses as they relate to designing and reporting of clinical trials. The limited success in bringing forward new therapies in these areas is likely proof of that premise. The ability to use preclinical and phase I and II trial data to predict which patients and which dosing regimens are more likely to benefit is perhaps the greatest challenge. Animal models continue to be refined in attempts to more accurately reproduce human sepsis and acute lung injury. Oncology research should serve as a model for optimizing the integration of pharmacodynamics and pharmacogenetics into trial design. The European Organization for Research and Treatment of Cancer provides a valuable template for nonfunded multicenter clinical trial success. The marked heterogeneity of the patient population and small signal (tested therapy)–to-noise (comorbidities) ratio makes identification of treatment effect difficult. Dedicated investigators still enroll ineligible patients who are included in intent to treat analysis. High enrolling centers create less problems in an adequate test of a new therapy. Much has been learned from negative trials as to value of post hoc subgroup and interim analyses. Debate continues on fair and appropriate end point of trials. Extrapolation of adult positive trial results to children is problematic. Conflict of interest issues which rested dormant for years are now at the forefront of discussion, and journal editorial board responsibility in this area is being recognized. Protocols may also help reduce heterogeneity of treatment across centers in clinical trials. This article reviews many of the problems encountered in clinical trial design and reporting and offers a perspective on dealing with them to the betterment of a clinical trial.

© 2008 Elsevier Inc. All rights reserved.

⁎Corresponding author.
E-mail addresses: dellinger-phil@cooperhealth.edu, cinel-ismail@cooperhealth.edu (R.P. Dellinger).
1. Introduction

Design and reporting of clinical trials in severe sepsis and acute lung injury is unique in comparison with most clinical research in that the patients are, by definition, critically ill with much heterogeneity and significant individual variability in risk of death. The character of these 2 patient populations poses particular important issues for trial design. Most of our current approach to clinical trial design and reporting in severe sepsis and acute lung injury comes from building progressively on lessons learned from previous trials. This manuscript discusses criteria for proceeding to clinical trials, conduct of clinical trials, options for clinical trial design, clinical trial reporting, and translation of clinical trial results to bedside care.

2. Challenges of clinical trials in sepsis and acute lung injury

Why are clinical trials in sepsis and acute lung injury so difficult? A primary reason is that novel agents in these conditions are characterized by a small signal-to-noise ratio for the treatment effect in a study population at risk for death from many other conditions [1]. It is quite possible that previous trials with experimental agents failed to show significant outcome benefits despite actually working, being overwhelmed by other events that masked the drug’s efficacy. Inexact definitions applied to heterogenous patients in diverse therapeutic environments limit the application of randomized controlled trial methodology to sepsis and acute lung injury (ALI). To overcome some of these barriers, it is important to use conservative sample size estimates, study a highly active drug, and choose a carefully defined clinical trial population. Recent evidence suggests that ALI is associated with systemic activation of inflammation and coagulation, that the degree if this activation correlates with outcome, and that low tidal volume ventilation ameliorates this activation point to the importance of lung protective strategies in clinical research in ALI [2-6]. New definitions that incorporate prognostic measures and reduce patient heterogeneity should allow more efficient patient enrollment [7]. Acquiring the ability to personalize subject enrollment in a fashion that identifies patients who have biology that matches the proposed treatment is likely to be a major step forward in producing positive clinical trials, for example, enrolling septic patients with altered sublingual flow to receive a therapy targeting improvement in flow or administering antiendotoxin therapy to a patient testing positive for endotoxemia. The inability to enroll suitable patients that allow at least an opportunity for the new therapy to work has major negative effect on power calculation [8]. Clinical coordinating centers may assist in limiting some of the variables intrinsic to human sepsis studies and ensuring that the subject enrolled meets the true intent of the protocol [9].

3. Pre-clinical and phase I to II trials in severe sepsis research

“Except on few occasions, the patient appears to die from the body’s response to infection rather than from it.” Sir William Osler (1904), The Evolution of Modern Medicine

For more than a century, the control of the systemic response to infection has been the target for research in sepsis. Why have so many past phase III clinical trials failed? [10] Is the problem with our preclinical data? Indeed, our animal models are not ideal for predicting performance of therapies in humans, and this problem remains a major impediment in translational research for sepsis [11].

We must accept that although existing animal models provide useful information, they are highly contrived and specifically designed to limit the number of variables that affect outcome, other than the direct intervention under study. The physiology, genetics, immunology and host response to infection differs significantly from one species to another, and with the exception of chimpanzees, substantially differs from humans [12]. Animal studies focus on previously healthy, young animals with a defined genetic background and without underlying diseases. They are given a complete diet before the experimental septic insult. Timing and magnitude of injury and the nature of the causative pathogen are often carefully controlled by the investigator.

The clinical situation is strikingly different in the intensive care unit (ICU) sepsis and acute lung injury patient populations who manifest an extremely variable genetic makeup, and are typically malnourished, old and immunocompromised. A multitude of underlying diseases exist in most study patients and the causative microorganism(s) is highly variable. The onset and pace of illness in humans is often poorly defined. These fundamental differences between animal models and actual patients make extrapolations on dosing and efficacy of agents based on preclinical studies imprecise and inaccurate.

Phase I-II clinical trials in sepsis and acute lung injury also have limitations as to the extent to which they can provide guidance on dosing, timing and duration of therapy for pivotal phase III studies [9]. Dosing and duration of experimental agents should ideally be derived from phase I to II studies using therapeutically relevant end points or highly predictive surrogate markers for efficacy. Since surrogates for efficacy are the perturbations of severe sepsis and sepsis-induced organ dysfunctions are also potential treatment toxicities, a control arm is useful. Regrettably, the only current generally accepted clinical end point for phase
III clinical trials in sepsis is improved mortality rates at 28 days. Septic patients have widely disparate mortality rates from a multitude of underlying diseases, multiple pathogens, and imprecise definitions of sepsis. These variables conspire to make accurate predictions on efficacy from limited phase II trials difficult at best. Moreover, there is no consistent surrogate marker that accurately predicts outcome of severe sepsis. However, phase II trials should be designed to evaluate the composite primary and secondary clinical end points as well as the biologic end points without excessive reliance on the primary outcome.

4. Randomized clinical trials

4.1. Identifying the best patient population for enrollment

What makes a randomized controlled trial in critically ill patients different from other randomized controlled trials? In contrast to noncritical care populations, screening of patients for eligibility is defined typically by service (critical care) or location of care (ICU) and not initially by presence of a specific disease (as opposed to a diagnosis of a certain type of cancer). Care given to the typical ICU patient is very complex and delivered by multiple different disciplines. Trial enrollment is typically determined by very strict inclusion and exclusion criteria, whereas if the trial is successful, the “effectiveness” of the drug is determined by application in more usual practice conditions.

Patients with severe sepsis are heterogeneous, and their outcome is often determined by factors totally unrelated to the sepsis condition. The definitions of sepsis and severe sepsis have been debated, and the difficulty in defining the study population is echoed by demonstration that dedicated investigators still produce significant enrollment of ineligible patients which are then included in the intent to treat analysis [13]. When a clinical evaluation committee is used to evaluate all case report forms to make an independent decision on whether the patient had been enrolled according to proper inclusion and exclusion criteria, a better treatment effect in the clinical evaluation committee adjudicated group than in the intention to treat group has been demonstrated [13]. Targeting sepsis or acute lung injury in the more severely ill increases the chance to show mortality benefit with a fixed sample size but the trick is to exclude those subjects that have a high likelihood of dying regardless of therapy. Rapid bedside tests to identify biochemical profiles that can be matched to targeted therapy would be ideal to enhance proper subject selection. The ability to define polymorphisms associated with specific genotypes by using rapid genetic testing may useful in the identification of patients more likely to develop certain types of infectious conditions and to potentially benefit from specific therapies [14].

4.2. Center effects on study outcomes (high vs low enrollers and quality of supportive care)

Almost all phase III trials in severe sepsis patients have failed despite promising results from smaller phase II trials. It is likely that the quality of sepsis therapy as well as the experience of study site clinicians influences the effectiveness of new study drugs. Besides patient heterogeneity, wide variations in quality and practice of standard care may have contributed considerably to the failure of large multicenter sepsis trials. There is evidence that change in critical care physician staffing alone may result in mortality reductions from 74% to 57% in patients with septic shock [15]. Favorable trends for a tested therapy may significantly increase when patients are excluded from the intent-to-treat (ITT) analysis who either did not fulfill the inclusion or exclusion criteria or did not receive adequate standard care (ie, adequate antimicrobial therapy, proper surgical source control, or inadequate supportive care). For example, in one international trial (International Sepsis Trial) of anti–tumor necrosis factor therapy in severe sepsis, the exclusion of patients with such confounding events resulted in an increase in relative risk reduction from 14.5% (ITT) to 26.5% as defined by a blinded scientific extramural review committee group [13].

4.3. Alternatives to randomized trials

Randomized clinical trials can also be confusing to the practitioner such as when one study of transfusion in critically ill patients demonstrates increased mortality and another study using transfusion as part of the therapy demonstrates that it is lifesaving [16,17]. Regardless of the result of clinical trials, there may still be reluctance by the bedside clinician to use a therapy. For example, meta-analysis suggesting increased survival with selected decontamination of the gut in select patient populations has not been assimilated into general clinical practice, likely due to fear of emergence of bacterial resistance [18,19]. Conflicts in meta-analysis concerning best therapy for prophylaxis of stress ulcers in severe sepsis may further confuse the practitioner [20,21]. Randomized controlled trial requirements cause divorce between care and research, whereas cohort, observational, or physiologic (noninvasive) studies may link care and research. Well-designed observational studies can offer similar estimate of magnitude of a treatment effect as a randomized control trial and will remain an important resource for decision-making.

4.4. Outcome measures

“Far better an approximate answer to the right question, which is often vague, than an exact answer to the wrong question, which can always be made precise.” John W. Tukey [22]
Although mortality end points for therapies targeting prevention of severe sepsis are difficult to achieve, they are typically expected for treatment studies. Severe sepsis mortality ranges from 20% to 50%, with most studies now demonstrating approximately 30%. Studies of septic shock demonstrate higher mortalities, from 40% to 75%. Acute lung injury and acute respiratory distress syndrome (ARDS) mortality range from 20% to 40%. The traditional primary outcome variable has been all-cause mortality 28 days after randomization, although later mortality comparisons are becoming more common and more important as follow-up of patients is extended. This is traditional in critically ill patients where mortality is expected to be greater than 10% over the first 28 days. Clinically relevant considerations as to the appropriateness of this end point for ARDS include the fact that 90% of patients who die after acquiring severe sepsis or ARDS die within 2 to 4 weeks. Choosing shorter time points are unlikely to demonstrate full benefit, and choosing longer time points make drug effect difficult to distinguish from other coexisting morbidities influencing outcome. When studies are carried out in sepsis groups with low mortality, potential alternatives to mortality end points include prevention of organ failure, prevention of progression of organ failure, prevention of progression to severe sepsis, and prevention of progression to ARDS. Because, other than mortality, the next most important effects are avoidance of organ failure, support end points of organ failure and support free days over time are also valid (ventilator-free days, renal replacement-free days). In these cases, it is important to standardize criteria for initiation and discontinuation of therapy across participating centers. Length of ICU and hospital stay are also important. Secondary outcome variables in confirmatory studies are important to offer likely reasons for primary outcome benefit (Table 1). Special populations of consideration as they relate to primary outcome variables include the elderly, where age is a very high-risk factor for mortality in both ARDS and sepsis [23,24].

An end point is defined as a “measure of the effect of an intervention on outcome (eg, success or failure) in a clinical trial in treatment or preventing a disease.” End points measured should be clinically relevant to the disease and question. The measurement process should ascertain the differences between therapies, if they exist, to distinguish effective from ineffective drugs. Timing of the measurement should be clinically relevant to the disease. Magnitude of the effect should correlate with clinical outcomes that are meaningful to the patient. How the analysis is performed may have significant effect on conclusions. The clinical end point is a direct measurement of how a patient feels functions or survives. As mentioned earlier, mortality has been the traditional severe sepsis trial end point, although resolution of symptoms of disease to allow discontinuation of organ support has validity in ALI/ARDS, that is, days alive and off mechanical ventilation. Admittedly, this end point is a blend of mortality and resolution of organ failure. A surrogate end point is a laboratory measurement or physical sign that is used for a substitute for clinical end point such as culture results, radiological testing, or histology. Such a definition of surrogate end point does not confer direct clinical benefit to the patient. In the pure sense, a surrogate end point is, for example, substituting resolution of hypotension or resolution of elevated cytokine levels for clinical outcomes of disease. Surrogate end points are likely very useful in early drug development (phase I studies) as proof of principle of drug effect and selecting candidate drugs. An example of successful surrogates in other diseases include lowering cholesterol in hopes of preventing cardiovascular disease, lowering blood pressure in hopes of preventing cardiovascular disease, and suppression of HIV viral load in hopes of treatment of HIV/AIDS. One potential problem with such surrogate end points is that toxicity, which may counterbalance beneficial effect, is not factored into this equation.

Surrogate end points in sepsis and acute lung injury must not only correlate with clinical outcomes but also take into account methodology and uncertainties in measurement, provide consistent results across product lines of a drug class, and evaluate and consider unmeasured harms and benefits.

### Table 1
Secondary outcome variables in confirmatory (phase III) trials

<table>
<thead>
<tr>
<th>Short term:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>● Days alive and off mechanical ventilation</td>
<td></td>
</tr>
<tr>
<td>● Days in ICU</td>
<td></td>
</tr>
<tr>
<td>● Days on pressor support</td>
<td></td>
</tr>
<tr>
<td>● Frequency of complications (neck/thorax, GI tract, barotraumas, hemodynamic)</td>
<td></td>
</tr>
<tr>
<td>● Organ function (eg, SOFA)</td>
<td></td>
</tr>
<tr>
<td>● Time to development/resolution of organ dysfunction</td>
<td></td>
</tr>
<tr>
<td>● PaO₂/Fio2 ratio</td>
<td></td>
</tr>
<tr>
<td>● Dynamic markers of activity—inflammation and hemostasis</td>
<td></td>
</tr>
<tr>
<td>● Morbidity assessment (eg, number and nature of ARDS)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Long-term (6 mo):</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>● All-cause mortality in hospital</td>
<td></td>
</tr>
<tr>
<td>● All-cause mortality at 3, 6, and 12 mo</td>
<td></td>
</tr>
<tr>
<td>● Quality of life</td>
<td></td>
</tr>
<tr>
<td>● Respiratory function</td>
<td></td>
</tr>
<tr>
<td>● Psychodynamic, renal, hepatic, and central nervous system function</td>
<td></td>
</tr>
<tr>
<td>● Morbidity assessment</td>
<td></td>
</tr>
</tbody>
</table>

### 5. Applying clinical trial results to clinical practice in the pediatric ICU

Trials done solely in pediatric patients support optimal pediatric clinical decision making. When common disease states exist for both pediatric and adult patient populations, however, such as in severe sepsis and acute lung injury, attempts are usually made to extrapolate from adult to
pediatric implications in some fashion. An example of how difficult it is to replicate and extrapolate findings from clinical trials in critically ill adults to critically ill children is reflected in the use of activated protein C in severe sepsis and surfactant therapy of acute lung injury.

After approval of recombinant activated protein C (rhAPC) in adults with severe sepsis, there was sporadic off-label use in pediatric patients with septic shock and multiple organ failure. A study was performed in septic children requiring vasopressors and mechanical ventilation powered to detect a 1-day difference in duration of organ failure. The study was stopped early for futility. Mortality was the same with a numerical increase and rate of intracranial hemorrhage in the treatment group, and the company has recommended that the drug not be used in children [25]. It should be noted that following the Recombinant human activated protein C worldwide evaluation (PROWESS) trial, which supported benefit of rhAPC in severe sepsis, a subsequent trial in less severely ill septic patients did not demonstrate survival benefit of rhAPC [26,27]. A new trial mandated by the European Regulatory Authority is being done in patients with septic shock [28].

Following several negative trials for surfactant in adult acute lung injury, the Pediatric Acute Lung Injury and Sepsis Investigators consortium performed a trial of surfactant delivered endotracheally, which demonstrated significantly greater mortality in placebo group with no difference in ventilator-free days [29].

5.1. Tidal volumes in acute lung injury

It would be difficult to replicate the National Heart Lung and Blood Institute ARD Snetwork low-tidal-volume strategy trial [30] in children, which showed improved outcome in adults, because measurement of tidal volume in young children is problematic owing to use of uncuffed endotracheal tubes. In addition, use of high-frequency ventilation is common, and extracorporeal membrane oxygenation (ECMO) is used for rescue therapy, making a further-limiting trial design.

In summary, clinicians must be wary in extrapolating the results of clinical trials in adults to the care of critically ill young children. This awareness has led to an increase in sponsorship of pediatric clinical trials [31]. Replication of adult trials in children remains challenging, however, because of the issues identified above.

6. Lessons learned from cancer research: European Organization for Research and Treatment of Cancer

Oncology research is an excellent model for research in severe sepsis and acute lung injury as it has been in place on national and international fronts much longer than sepsis/acute lung injury research and has become very mature.

7. Editing and review issues for clinical trial reporting and conflict-of-interest issues

Outcome reporting bias is a significant clinical research problem. Unreported outcomes are a particular problem. It is not unusual to have outcomes of clinical interest described in methods but not presented in results. An evaluation of 519
Conflict of interest is appropriately becoming a more and more important factor in our scrutiny of clinical trials. Association between competing interests and author conclusions in clinical trials has been reported [37]. In addition, a significant number of biomedical researchers have dual university and biotechnology firm affiliation, and high percentage of academic research is now supported by biotechnology firms from 46% in 1986 to 92% in 1996 [38]. Institutions often hold equity in businesses engaged with research at the same institution, with a minority but significant portion being startups. More concerning is that policies and disclosures of conflict of interest were highly variable across institutions and journals. A study in 2003 showed that industry studies are significantly more likely to induce favorable outcome for interventions than nonindustry studies [39]. Possibilities for this occurrence include industry willingness to fund only therapies more likely to be winners, study design that favors experimental therapy, and publication bias in multiple reporting of positive results. Transparent reporting of funding as well as disclosure of intellectual bias is essential to allow readers to make an informed judgment about study reports as well as editorials, opinion pieces, and clinical perspective articles. Intellectual bias (advocate of a particular therapy or position based on previous research, previous publications, or previous interactions with a therapy) is a poorly recognized but important factor that needs to be disclosed.

8. Conclusions

Criteria for preclinical support to move into clinical trials remains tenuous at best and is determined by discussions between trial sponsor and the regulatory agency section assigned to that study. Significant variability from study to study remains. The value of clinical evidence other then randomized clinical trials is now more appreciated. Significant variability may exist in clinical management among participating investigative centers. At each center, study protocol deviations decrease over time. Both of these findings handicap achieving a positive trial outcome, even with a truly beneficial therapy. The most acceptable outcome measure remains to be mortality, although innovative alternatives are now being proposed and need to be evaluated. Pediatric patients with severe sepsis and acute lung injury pose particular problems in translating adult trial data to pediatric use. Conduct of the data safety and monitoring brand is an important aspect of clinical trial design. Many lessons are to be learned from international cancer research which precedes sepsis and ALI research by a significant number of years. The job of the journal editor is important in reporting trial results, particularly as it relates to complete data reporting and conflict of interest issues. In the end the goal of proper design, enactment and reporting of

---

**Table 3** Outcomes and analysis score

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>All-cause mortality</td>
<td>3</td>
</tr>
<tr>
<td>Hospital/ICU length of stay</td>
<td>2</td>
</tr>
<tr>
<td>Surrogate physiologic end point</td>
<td>2</td>
</tr>
<tr>
<td>Complications/resources use</td>
<td>1</td>
</tr>
<tr>
<td>Validated organ dysfunction score applied to all groups</td>
<td>1</td>
</tr>
<tr>
<td>ITT on mortality</td>
<td>2</td>
</tr>
<tr>
<td>ITT on other end points</td>
<td>2</td>
</tr>
<tr>
<td>Sample size calculation</td>
<td>2</td>
</tr>
<tr>
<td>Presentation of data allowing replication of analysis</td>
<td>1</td>
</tr>
<tr>
<td>CI reported</td>
<td>1</td>
</tr>
<tr>
<td>Multivariate analysis to adjust for imbalances</td>
<td>1</td>
</tr>
</tbody>
</table>

ITT, intent; CI, confidence intervals.

trials reported in 555 publications in 2000 revealed that 75% and 64% of efficacy and harm outcomes, respectively, were not fully reported [33]. Reported outcomes for which survey responders deemed clinically important relative to the clinical study are often unreported [33]. The most recognized reasons for omitting one or more outcomes per trial in rank order includes space constraints, not clinically important, not statistically significant, not yet submitted and not yet analyzed [33]. Where space constraints were listed as the reason for not reporting, this was identified as journal-imposed in one third and author-imposed in two thirds. Problems created by reporting bias include detrimental effect on critical appraisal of individual trial results and potential for favoring use of harmful interventions. Potential ways to address and avoid this problem include (1) trial registration before completion, (2) protocols available in the public domain, and (3) transparency of analytic plan and selection of reported outcomes [34]. A scoring system for methodologic quality outcomes and analysis reporting has been described by Graf et al [35] and is demonstrated in Tables 2 and 3. Using this scoring system, some improvement in quality of sepsis trials was demonstrated between 1976 and 1998, rising from 20 in 1976 to 35 in 1998 (out of a maximum score of 57). It should also be recognized that a “poor-quality” score may reflect poor description and reporting in the publication and not necessarily poor quality of the study.

What about the responsibility of the editor and reviewers to assure quality reporting? Editor and reviewer responsibilities include reviewing discrepancies between methods and results sections and ensuring that all prespecified and clinically relevant outcomes are being reported adequately and completely independent of results. There should be support in the submitted manuscript for ethical conduct of the trial [36]. To limit outcome reporting bias, researchers and journal editors should ensure that complete data be provided for all prespecified trial outcomes [33]. The increasing use of journal internet sites should help to alleviate concerns over space restrictions.
clinical trials is to insure and that positive results from clinical trials are likely to have a clinical effect and that promising therapies are not inadvertently discarded.

References
