Published in final edited form as:
Crit Care Resusc. 2013 December ; 15(4): 301–310.

Protocolized Care for Early Septic Shock (ProCESS) statistical analysis plan

Francis Pike[1,2] **[Assistant Professor, and Director of Biostatistics and Data Management Core]**, **Donald M Yealy**[3] **[Professor and Chair]**, **John A Kellum**[2] **[Professor and Vice Chair for Research]**, **David T Huang**[2,3] **[Associate Professor]**, **Amber E Barnato**[4] **[Associate Professor of Medicine, Clinical and Translational Science, and Health Policy]**, **Tammy L Eaton**[2] **[Project Manager]**, **Derek C Angus**[2] **[Professor and Chair, and Director]**, and **Lisa A Weissfeld**[1,2] **[Professor] on behalf of the ProCESS Investigators**

[1]Department of Biostatistics, Pitt Public Health, University of Pittsburgh, Pittsburgh, Pa, United States

[2]Department of Critical Care Medicine, Clinical Research, Investigation, and Systems Modeling of Acute Illness (CRISMA) Center, University of Pittsburgh, Pittsburgh, Pa, United States

[3]Department of Emergency Medicine, University of Pittsburgh, Pittsburgh, Pa, United States

[4]Department of Medicine, University of Pittsburgh, Pittsburgh, Pa, United States

Abstract

Background—The Protocolized Care for Early Septic Shock study is a randomised, multicentre, prospective, three-arm, parallel-group trial of alternative resuscitation strategies for early septic shock.

Objective—To state our analysis plan for trial data.

Methods—Our plan is to guide data collection and analysis using pre-existing definitions and testing, with local consensus-based efforts where needed. We examine protocolised care (two experimental approaches) and compare this to usual “wild type” care.

Results—Our plan is to address three aims (clinical efficacy, biology of illness and recovery, and costs and cost-effectiveness) and four hypotheses, and we specify rules for handling data and determining outcomes.

Conclusion—By using measures to maintain study conduct and analysis rigour, we hope to improve understanding of early septic shock resuscitation and care of patients.

*Correspondence:*lweis@pitt.edu.

Competing interests
None declared.

Overview

Background and goals

In 2001, Rivers and colleagues published a seminal manuscript on the early resuscitation of patients with septic shock.[1] They observed a marked improvement in short-term mortality when using a structured, physiological approach to resuscitation in the first 6 hours of care, delivered via a protocol of fluids, vasopressors, blood or inotropes. The absolute change in mortality was high (46.5% in control patients, compared with 30% in protocol patients), but use of the approach and protocol specified by Rivers and colleagues is hampered by concerns about its generalisability and the contribution of individual components.[2]

The Protocolized Care for Early Septic Shock (ProCESS) study is a randomised, multicentre, prospective, three-arm, parallel-group trial of alternative resuscitation strategies for early septic shock. Institutional review board (IRB) approval was obtained from the University of Pittsburgh and all participating sites, and the trial is registered with Clinical Trials.gov (NCT00510835).

Our primary goal is to determine the clinical efficacy of two protocolised resuscitation strategies, compared with usual care. We will also assess the effect of these resuscitation strategies on markers of biological pathways and on cost and resource use. Our design randomises patients to receive one of two resuscitation strategies or usual care ("wild type", without any structured care). The experimental resuscitation strategies are:

- early goal-directed therapy (EGDT), based on the Rivers protocol and guided by systolic blood pressure, central venous pressure and central venous oximetry;[1] and
- protocolised standard care (PSC), an approach that delivers fluids and vasopressors based on simple bedside criteria without the use of invasive monitoring.

Our trial is harmonised with but independent from similar studies in Australia and the United Kingdom.[3] The three studies target the same group and use the same basic approach — resuscitation in the first 6 hours of recognition of septic shock, testing the River's approach in one arm, and using a randomised, controlled design. The leaders of the three trials are maximising the consistency of their data collection to allow a future patient-level meta-analysis, allowing more insight into early septic shock care. A separate study group drawn from the parent trials will define research questions and an analysis plan before the data merge.

Here we report our statistical analysis plan for the ProCESS study, before unblinding of researchers and perarm outcome assessments.

Patient population

Based on census data from the emergency departments (EDs) of participating sites and United States federal claims data, we are enrolling a population that we expect to be 56% men, 68% white, 25% African American and 7% other races. We expect this distribution but are enrolling patients without regard to sex, race or age. The trial is not designed to make inferences about non-black minority groups.

Inclusion criteria

Our study is using similar inclusion criteria to that of Rivers and colleagues.[1] Patients must:

- be ≥ 18 years of age
- have a suspected infection
- meet two or more of the criteria for systemic inflammatory response syndrome, and
- have refractory hypotension (systolic blood pressure < 90 mmHg despite an intravenous [IV] fluid challenge of ≥ 1000 mL over a 30-minute period), or evidence of hypoperfusion (blood lactate concentration > 4 mmol/L). To identify refractory hypotension, we initially required a 20 mL/kg minimum crystalloid bolus over 30 minutes (identical to that of Rivers and colleagues[1]) but modified this to the simpler 1000 mL bolus in April 2010 to ease logistics.

Exclusion criteria

We are excluding patients who:

- are currently pregnant
- have a primary diagnosis of acute cerebral vascular event, acute coronary syndrome, acute pulmonary oedema, status asthmaticus, major cardiac arrhythmia, active gastrointestinal haemorrhage, seizure, drug overdose, burn or trauma
- need immediate surgery
- have a CD4 count < 50/μL
- have an absolute neutrophil count < 500/μL
- have a “do not resuscitate” code status or an advance directive restricting implementation of the protocol
- have a contraindication to central venous catheterisation
- are likely to refuse a blood transfusion (eg, Jehovah’s Witnesses)
- have a treating doctor who deems aggressive care unsuitable
- are participating in another interventional study
- have been transferred from another inhospital setting.

Objectives and aims

Aim 1—To compare the clinical efficacy of alternative resuscitation strategies for patients with septic shock, using sequential hypothesis testing, we are testing the following:

- Hypothesis Ia: structured care (EGDT and PSC) will produce superior short-term mortality outcomes compared with usual care.
- Hypothesis Ib: if Ia is true, EGDT will produce superior outcomes to PSC.

Aim 2—To understand the mechanisms of illness and recovery and how resuscitation strategies affect them, and affect clinical outcomes, we are testing the following:

- Hypothesis IIa: protocolised resuscitation changes the expression of markers of illness and recovery.
- Hypothesis IIb: the clinical efficacy of protocolised resuscitation changes markers of illness and recovery.

Aim 3—We aim to assess the costs and cost-effectiveness of the alternative resuscitation strategies.

Variable definitions

Primary outcome

The primary outcome is all-cause hospital mortality, truncated at 60 days, which is parallel to the approach of the Acute Respiratory Distress Syndrome Network trials of the National Heart, Lung, and Blood Institute. We will unblind researchers and begin analyses only after all data collection forms are complete, to the best of the abilities of the sites, and at least 120 days have passed from the last enrollee entering the trial.

Secondary outcomes

We will assess survival at 90 days and 1 year, clinical evidence of organ dysfunction and, in subsets of patients, absolute values and changes in markers of inflammation, oxidative stress, cellular hypoxia, coagulation and thrombosis. As part of Aim 3, we will also assess inpatient resource use, up to 60 days, including duration of mechanical ventilation, acute dialysis, hospital stay, intensive care unit stay and, in subsets of patients, total hospital charges. We will assess return-to-work, usual activities and health utility using EuroQol-5D scores at 90 days.

Safety

At trial design and start, we expected this study population to have an inhospital mortality rate in the control arm of 30%–46%, based on existing data.[4] To ensure optimal patient exposure and safety, the granting agency, independent safety board and coordinating centre is tracking overall mortality throughout the trial. Only the data and safety monitoring board has seen any per-arm outcome data.

Our reporting plan maximises the ability to detect any signal of differential treatment-related event rates across study arms without being encumbered by large numbers of reported events that accompany the illness (eg, background events). Therefore, our plan is to collect:

- detailed information regarding all serious adverse events occurring until Hour 72
- central venous oximetry catheter serious adverse events for the duration of hospitalisation or until Day 60, whichever period is the shorter
- all late-occurring (after Hour 72) serious adverse events detected by sites and potentially related to the study intervention (including late infections).

Analysis principles

- All primary analyses will be conducted on an intention-to-treat basis.
- Only patients who decline use of outcome data will be excluded from analysis. Exclusions will be reported per arm (see below). Patients who have protocol violations are analysed per the assigned treatment arm.
- All hypothesis tests will be two sided, with an α of 0.0494 unless otherwise specified.
- All analyses are unadjusted unless otherwise specified.
- Subgroup analyses are performed irrespective of treatment efficacy.

Design

Data collection and follow-up

There is the potential for two separate datasets: one for all randomised patients, and one with only randomised patients for whom data are available. For patients who decline participation by withholding consent, we will collect a limited set of demographic data (to the extent allowed by each site IRB) to compare patients who did and did not enrol in the study.

The stages of data collection and follow-up are randomisation, baseline, intervention (Hour 0–Hour 6), postintervention (Hour 7–Hour 72), other hospital follow-up (Day 2–Day 60 or at discharge), postdischarge survival, and data collection supporting Aims 2 and 3. The data collected at each of these stages are as follows.

Randomisation

- Patient demographics and inclusion and exclusion criteria.

Baseline

- Sociodemographics: age, sex, race, ethnicity, residence before admission and employment status.
- Comorbid conditions and medications: Charlson comorbidity index score, chronic illness components of the Acute Physiology and Chronic Health Evaluation (APACHE) III score and exposure in the preceding 7 days to antibiotics and other selected medications.
- Site and aetiology of infection: assignments will be made retrospectively, based on the criteria of the Centers for Disease Control and Prevention, and the Interscience Conference on Antimicrobial Agents and Chemotherapy, and using the schema adopted previously in several multicentre trials on severe sepsis.[5,6]
- Severity of illness: APACHE III, sequential organ failure assessment (SOFA) scoring systems.[6-9]

Intervention period (Hour 0–Hour 6)

- Time of randomisation: the time of randomisation is “time zero” for data collection, although data are collected before randomisation, to assess care.
- Measurements and therapies (all patients): we measure vital signs initially and hourly, until the end of Hour 6. Therapy data include IV fluid volumes, packed red blood cell transfusions, vasoactive agents and inotropic agents. Data are also collected on central line placement and mechanical ventilation before and during the protocol period. We are also recording data on all prehospital and prerandomisation IV fluid administration.
- Additional measurements (EGDT arm only): central venous pressure and central venous oxygen saturation hourly until the end of Hour 6.

Postintervention period (Hour 7–Hour 72)

- Haemodynamics and therapies: vital signs are recorded at Hour 12, Hour 24, Hour 48 and Hour 72.
- Organ dysfunction: we are collecting daily SOFA scores and will use the worst level recorded for each organ system over the time.

Other hospital follow-up data (Day 2–Day 60 or at discharge)

- Daily SOFA scores while in the ICU (data collection resumes for patients who are discharged from the ICU and readmitted within 48 hours of initial admission; any subsequent ICU discharge and readmission will not be collected).
- Hospital location (hospital floor or ICU), timing and type of cointerventions (eg, steroid supplementation and activated protein C) and adverse events.
- At hospital discharge, we are assessing and recording ongoing renal and respiratory support, and discharge disposition (to home, nursing home or rehabilitation facility, etc.). Site investigators will assign the source of original infection.

Postdischarge survival

- We are collecting 90-day mortality data, through a National Death Index (NDI) search or, for more recently enrolled subjects, by direct contact with the patient or their listed contacts.
- Survivors enrolled after the protocol modification approved in May 2011 by the local IRB will be asked (by direct contact) to complete the EuroQol-5D[10] and questions about their return to work and usual activities at 90 days.
- We are collecting long-term survival status (through to 1 year after discharge) through the NDI. There is a 2-year time lag before these data are available, so for patients enrolled near the end of the trial, we will perform primary analyses before their 1-year follow-up data are available, and a shorter follow-up will be handled with censoring.

Data collection for Aim 2 and Aim 3

- Aim 2: we are collecting blood and urine samples to analyse selected biomarkers. For blood sampling, we are collecting 30–35 mL of blood at four times (Hour 0, Hour 6, Hour 24 and Hour 72) with maximum of 140 mL of blood drawn over the study. For urine sampling, the ProCESS biorepository is also collecting urine samples from the ProCESS lab cohort at Hour 0, Hour 6, Hour 24 and Hour 72.
- Aim 3: we are collecting resource-use data to analyse costs and cost-effectiveness. The primary source of this information is the patient's data collection form and, for a subset of patients, the patient's UB-04 form (the National Uniform Billing Committee institutional provider bill).

Treatment allocation

Assuming entry criteria are met, each patient receives a study identification number and treatment allocation at enrolment. We randomise at a 1:1:1 ratio in variable blocks at each institution.

Power and sample size

We initially designed the study to have 80% power to detect an absolute risk reduction of 6%–7%, at a nominal significance level of 0.05, based on an expected overall event rate of 30%–46%. This required an equal allocation of 650 patients into each of the three study arms for a total sample size of 1950. We planned two interim analyses, when one-third and two-thirds of total enrolment were reached. We completed the first interim analysis with no recommendations for change in data or safety monitoring.

During the trial, we observed an overall blinded event rate of about 20%, the same as occurred in a recent trial with a similar design.[11] After consultation with the DSMB and the federal funding agency in February 2013, we resized the trial to a new total target of 1350 patients, to account for the lower observed overall event rate. The resizing retained the same pretrial targeted 80% power to detect an absolute risk reduction of 6%–7%. After "spending" about 0.0005 α during the first interim analysis at 650 patients, and after resizing to 1350 (which removed the requirement for a second interim analysis), the α required for our hypothesis tests at the close of enrolment will be 0.0494.

For the Aim 2 hypotheses, group sample sizes of 400 and 200 for the combined protocolised arms and control arms, respectively, allow 80% power to detect mean cytokine differences in the range 0.12–2.5 units. This assumes an exchangeable correlation structure, autocorrelation varying from −0.8 to 0.8, an SD of a single measurement varying from 1–10 units, four time points and a significance level of 0.05.

Consent

All patients or their legally authorised representatives provided consent for trial participation, with each site following the local regulations. We considered but did not use alternative consent strategies, including an exception or waiver process.

Permanent discontinuation

Patients who initially consent but later withdraw consent will be asked to agree to allow current data to be used for analysis. If they agree for their existing data to be used, they will be included and analysed on an intention-to-treat basis.

Statistical analysis

Trial profile

We plan a Consolidated Standards of Reporting Trials diagram to detail the movement of patients through the study. This diagram will include total patients screened, number who met inclusion and exclusion criteria, and number included in the study.

Baseline comparisons and assessment of randomisation

To assess randomisation success, we plan to tabulate the distribution of baseline variables across the study arms, and to summarise discrete variables by frequencies and percentages. We will report continuous variables as either means with SDs or as medians with interquartile ranges.

Process measures and concomitant treatments

Process measures—We will assess adherence to the experimental protocols at Hour 2, Hour 4 and Hour 6, based on prespecified actions and goal achievements. Hour 6 will be used for overall adherence analysis, but we will evaluate earlier and sustained adherence and the relationship to outcomes as a secondary effort.

Concomitant treatments—We will track ancillary care during Hour 0–Hour 72, including delivery and timing of antibiotics, activated protein C, steroids and other vasoactive agents.

Treatment limitation—Only counts and percentages will be reported for patients for whom there was a limitation or withholding of treatment. This refers to a bedside doctor withdrawing or withholding treatment that might otherwise prolong life if the treatment is no longer considered appropriate for that patient.

Consent and discontinuation of study treatment—Only counts and percentages will be reported for patient consent and permanent discontinuation of treatment.

Primary outcome: analysis of Hypothesis Ia and Hypothesis Ib

Unadjusted test of treatment effect—We will test the hypothesis that protocolised resuscitation is superior to usual care by comparing the difference in mortality proportions in the combined EGDT and PSC arms versus usual care. We will use an unadjusted two-sample test of proportions with an interim analysis adjusted P of 0.0494. If the null hypothesis is rejected, we will test the difference in mortality proportions between experimental arms using the same method. We will test all other treatment arm comparisons as secondary analyses. We also plan exploratory subanalyses in which the first patients at each site and arm are excluded to examine if a “warm-up” effect exists.

Modelling to examine potentially confounding factors—We will also fit logistic regression models to adjust for independent variables that were deemed to be imbalanced after randomisation or of clinical importance. We will explore the main effects and interaction models via stepwise selection or penalised regression approach to arrive at the most parsimonious model with the best fit, as determined by the Hosmer–Lemeshow test.[12] Treatment effect will be expressed as an odds ratio with 95% confidence intervals.

Modelling to adjust for potential institution effects—To address site variation, we will fit generalised linear mixed models[13] with a site-specific random effect to capture this heterogeneity and to adjust for potential confounders as described above.

Secondary and tertiary outcomes analyses

We will test the hypothesis that protocolised care changes long-term survival, compared with usual care. If this hypothesis is affirmed, we will then test whether the EGDT arm is different from the PSC arm. We will plot Kaplan–Meier curves for aggregated and individual experimental arms, testing the equality of survival curves using a log-rank test with a *P* of 0.0494.Long-term survival data from the NDI

We will also construct multivariable Cox proportional hazards models,[14] include treatment assignment as an independent predictor, and adjust for baseline covariates using selection models similar to those described under Primary outcome: analysis of Hypothesis Ia and Hypothesis Ib (above). We will assess the goodness of fit for this model via residual analysis and tests of proportionality. If the proportional hazards assumption is not met, we will use Gray's[15] spline-based extension of the Cox model. This provides time-varying estimates of regression coefficients, allowing the hazard associated with a particular covariate to vary during follow-up. Once we have selected the most parsimonious model, the magnitude and significance of the hazard associated with the treatment covariate will be our estimate of treatment effect.

Non-mortal end points—We will examine non-mortal end points (eg, SOFA and other scores) with descriptive statistics and generalised linear or linear mixed models to account for the different nature of the outcomes and possible repeated measures. The descriptive statistics will include means and frequency distributions, followed by corresponding statistical tests (analysis of variance [ANOVA] for three-group tests, *t* tests for two-group tests, and non-parametric exact tests for categorical outcomes).

Markers of inflammation, oxidative stress, cellular hypoxia, and coagulation and thrombosis—In a subset of 600 patients (200 per arm), we will analyse markers of inflammation (tumour necrosis factor, interleukin [IL]-6, IL-10), oxidative stress (urine isoprostane), cellular hypoxia (lactate), and coagulation and thrombosis (D-dimer and thrombin–antithrombin III complexes). This subset will consist of 600 randomly chosen patients, 300 from the first half of the trial and 300 from the second half of the trial. We will also study biomarkers related to these mechanisms and to organ injury attributable to sepsis.

Analysis of Hypothesis IIa—We plan a descriptive analysis, a primary analysis and exploratory techniques to identify potential clusters of interest within Aim 2.

- Descriptive analyses include computation of mean values and SD at each time point. Correlation matrices of the behaviour of each marker over time will be computed. The validity of the sample will be examined through comparisons of the three treatment arms using ANOVA for continuous outcomes and χ^2 tests for categorical outcomes.
- Primary analysis consists of an application of statistical methods for the analysis of repeated measures using a mixed model or a generalised estimating equations[16] approach. These models include the marker as the outcome, and time, treatment group assignment, and a time-by-treatment group assignment term in the model. We will assume an exchangeable correlation structure for the analyses of these data. The test of the significance of the interaction terms in these models will provide a test of differences across time for the three treatment arms. Should missing data prove problematic, we will use methods that address this issue directly, such as a pattern mixture model[17] approach.
- Exploratory techniques to identify potential clusters of interest will be done by clustering trajectories into groups. This method, implemented using PROC TRAJ (SAS Institute), allows the number of desired clusters to be user-selected or analysis-selected. This procedure is quite general and includes continuous outcome data as well as truncated outcome data. A common problem when analysing changes in markers is truncation of values at the lower limit of detection. However, methods such as a normal model for truncated outcome data can be applied using PROC QLIM) (SAS Institute). This approach can be used for repeated-measures outcomes and for trajectory analyses.

Analysis of Hypothesis IIb—The analysis plan for this hypothesis is similar to the analysis plan for Hypothesis IIa, with one major difference: modelling includes clinical outcome (mortality or morbidity, such as SOFA score) as a predictor. Initially, models will include time, clinical outcome, and time-by-clinical outcome interactions. A test of the significance of the time-by-clinical outcome interaction term will indicate if the behaviour of the marker differs by clinical outcome over time. To explore how these changes occur with treatment, a series of models will be fitted, including as potential variables: time, treatment, clinical outcome, time-by-treatment interaction, clinical outcome-by-time interaction and clinical outcome-by-treatment interaction. Terms will be retained based on statistical significance and validity. The significance of the clinical outcome-by-treatment interaction term will indicate a difference in profiles across the treatment and mortality groups. To test this effect over time, a time-by-treatment-by-clinical outcome term can be included in the model.

Secondary analyses and analytical issues related to Aim 2—For some analyses, we will treat the marker as a covariate. For example, we can summarise the trajectory of a marker through an estimated slope, a maximal value across time points, a mean value across time points, or by growth curve modelling. Using any of these approaches, the summary variable can then be included in a linear, logistic or survival regression model. As a further

refinement, errors-in-variables models[18] can be used to account for measurement errors. Alternatively, a joint modelling approach can be used.

Analyses and analytical issues related to Aim 3—We will conduct this analysis following the principles and recommendations of the US Panel on Cost-Effectiveness in Health and Medicine[19] and a position statement of the American Thoracic Society.[20] We will use methods developed as part of our prior cost-effectiveness assessments of other interventions and monitoring tools for sepsis, shock and organ dysfunction.[21,22] We will measure "base case" incremental cost-effectiveness ratios (ICERs), expressed as hospital costs per hospital survivor, from the US hospital perspective, and ICER from the US societal perspective ("reference case"), expressed as lifetime costs per survivor, costs per life-year and costs per quality-adjusted life-year. Hospital costs will be determined by collecting information on resource use in the data collection form, and multiplying resources consumed by cost weights derived from a detailed external cost database. Quality of life will be determined, for a subset of patients, using 90-day EQ-5D data. Estimates of longer term costs and quality of life will be derived from published sources. These measures will be incorporated in a simulation model to produce the reference case.

The effect of adherence on the estimates of treatment effect—We will assess and report adherence measures that include several approaches: intention-to-treat (ITT), as-treated (AT), per-protocol (PP), and instrumental-variables approaches.[23]

The standard approach to assessing the treatment effect in the presence of potential non-adherence is the ITT approach, in which all patients are analysed as if they received the treatment as intended. The limitation of this approach is that the estimates of treatment effect are biased toward the null hypothesis, resulting in a potential underestimation of the treatment effect.

The AT approach can provide an upper boundary on the treatment effect when non-adherence is random. As it is likely that non-adherence in our trial will not be random, this approach will not be the focus of the analyses.[24]

The PP approach focuses on the analysis of patients who received the treatment as specified in the protocol. These estimates will be computed as a comparison with the estimates obtained using the ITT approach. The instrumental-variables approach will also be considered when estimating the treatment effect in the presence of potential non-adherence. This method is based on the computation of the complier-average causal effect, which measures the causal effect of the intervention on the patients who received it as intended by the original group allocation.[25,26]

Subgroup analyses—We will conduct prespecified subgroup analyses to understand the treatment effect, and to identify subgroups of patients for whom the treatment was particularly beneficial and/or harmful. These also allow future hypothesis generation. The subgroups are predefined to limit biases after unblinding.

We will first use descriptive testing for differences in treatment effects across the groups, through ANOVA techniques and through testing of interaction terms in statistical modelling. These interaction terms will consist of the interaction between the subgroup and treatment assignment. The subgroups of interest include (but are not limited to) the following variables: type of shock, source of infection, race, sex, age, anaemia status, timing of resuscitative actions and goal achievement, and level of adherence.

We will also apply newer methods to address bias when testing for treatment effects in subsets in clinical trials. One of these approaches, subpopulation treatment effect pattern plots, as outlined by Bonetti and Gelber,[27] creates plots of overlapping regions of potential covariates of interest. These plots provide a visual picture of the treatment effect over the full range of the covariate, and are easy to interpret, since they rely on traditional analysis methods for the plot presentation. Another approach is that outlined by Tibshirani and colleagues,[28,29] which relies heavily on the use of the least absolute shrinkage and selection operator for meaningful variable selection in interaction models.

Missing data

Missing data (due to incomplete forms or withdrawal from the study) are handled by the use of weighted estimating equations (in which the weights are functions of the probability of missing data) or by multiple imputation methods.[30-32] These are standard statistical approaches for the handling of missing data, that can be readily implemented in statistical software packages.

Conclusion

We describe, before any data unblinding, our approach to analysing the data from the ProCESS early resuscitation trial. We anticipate that this framework will enhance the utility of the reported result and allow readers to better judge the impact.

Acknowledgments

Funding

Our trial is supported by grants NIH/NIGMS P50 GM076659 and NIH/NHLBI K12 HL109068. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the National Institute of General Medical Sciences or the National Heart, Lung, and Blood Institute.

Appendix 1. Secondary clinical outcome definitions and rules for the Protocolised Care for Early Septic Shock study

Rules for resolution of shock

- Shock is resolved when all four criteria for resolution of shock (below) are met for 72 hours. The following provisions also apply:
  - ➢ Because of data collection limitations, eligibility for shock resolution will be within the first 7 days of study randomisation. Patients who do

not meet shock resolution criteria by Day 7 will be classified as unresolved at Day 7.

- ➢ The assessment for the resolution of shock begins at the first time point that the patient meets all four criteria simultaneously. Assessment can also begin when data are not obtained (see below).
- ➢ The resolution of shock is not achieved if the patient dies within the 72-hour assessment period.
- ➢ If a patient was discharged alive from the hospital, having met criteria for resolution but before completing the entire 72-hour assessment period for resolution of shock, the criterion is met at hospital discharge.
- ➢ Resetting the clock: if a patient, having met the criteria for shock resolution, fails to meet criteria at any time during the 72-hour assessment period, the clock (72-hour assessment) will restart once the patient meets all four criteria again.

1. Resolution of vasopressors, defined as:

- Dose of vasopressors is zero.
- No vasopressor information is entered (information is missing).
- None of the following is administered: dopamine, adrenaline, noradrenaline, vasopressin, phenylephrine.
- Dobutamine is not a vasopressor for this determination.
- Stand-alone analysis of resolution of vasopressors only (not to be used within criteria for resolution of shock):
  - ➢ Because of data collection limitations, eligibility for vasopressor resolution will be within the first 7 days of study randomisation. Patients who do not meet vasopressor resolution criteria by Day 7 will be classified as unresolved at Day 7.
  - ➢ A patient must survive for at least 24 hours after vasopressor use is not reported.
  - ➢ If the patient is discharged from the intensive care unit and survives ≥ 24 hours, the patient is to be free from vasopressor use.

2. Systolic blood pressure (SBP) ≥ 90 mmHg.

- ➢ Defined as: all SBP values entered are ≥ 90 mmHg, or no SBP is entered (information is missing).

3. Serum lactate < 4 mmol/L, defined as: all recorded values < 4 mmol/L or no values recorded.If SBP is < 90, the SBP criteria will be considered met if the patient has

been discharged from the ICU (with no readmissions within 48 hours) and death has not occurred.

4. Patient has not received > 4 L intravenous (IV) fluid within a 24-hour period, defined as: total IV fluids received in a 24-hour period is not > 4 L. Do not use multiplier (for colloids) and do not use any blood product volumes for these criteria. IV fluid totals are only recorded in the electronic data collection form until Hour 72 only. Therefore, after the initial 72 hours, this criterion is always presumed to be met.

Rules for duration of mechanical ventilation (MV)

- We define MV as invasive ventilation only; it does not include non-invasive techniques such as continuous positive airway pressure or bi-level positive airway pressure mask ventilation.
- Duration of MV is the number of consecutive* days that a patient requires invasive MV. *If a patient is free of MV for < 48 hours (or 2 days when hours are not recorded), these days are counted as part of the duration of MV.
- Patients known to require chronic invasive MV before hospital admission are not included in this category.
- Duration of MV is calculated for patients who start invasive MV within the first 7 days of study randomisation.
- To be considered free of MV, the patient must survive for at least 48 hours (or 2 days when hours are not recorded) after the last recorded use of MV.

Rules for renal replacement therapy (RRT)

- Types of RRT in this measurement include peritoneal RRT, continuous renal replacement therapy and haemodialysis.
- Duration of RRT is defined as the number of consecutive* days that a patient requires RRT. *If a patient is free of RRT for < 72 hours (or 3 days when hours are not recorded), these days are counted as part of the duration of RRT.
- Patients with a history of chronic dialysis prior to hospital admission are not to be included in this category.
- Duration of RRT is calculated for patients for whom RRT started within the first 7 days of study randomisation.
- Manual adjudication of RRT duration will be used for all patients who:
  - ➢ die in hospital after completing RRT therapy (ie, all patients who are free of RRT but do not survive to hospital discharge)
  - ➢ are discharged from the hospital alive within 72 hours (3 days) of last recorded RRT.

Rules for acute kidney injury (AKI) recovery

- We define AKI as any of the following (see modified risk, injury, failure; loss, end-stage renal disease [RIFLE] and Kidney Disease: Improving Global Outcomes [KDIGO]? criteria for additional staging in Appendix B):
  - ➢ increase in serum creatinine (SCr) level by > 0.3 mg/dL within any 48-hour period
  - ➢ increase in SCr level to > 1.5 times baseline level
  - ➢ urine volume < 0.5 mL/kg/hour for any 6-hour period.
- Full AKI recovery: all AKI patients; return to < 1.5 times baseline SCr level (no RIFLE criteria met)
- Partial AKI recovery: all AKI patients; return to a lower RIFLE criterion (eg, 3 to 2 or 1, or 2 to 1)
- Full AKI recovery after RRT, meets all of the following:
  - ➢ independence from RRT for > 72 hours (or 3 days when hours are not recorded)
  - ➢ documentation of at least one estimated glomerular filtration rate (eGFR) level > 30 mL/minute/1.73 m^2 (or no less than baseline level if baseline < 30 mL/minute/1.73 m^2 and SCr no more than 1.5 times baseline level within 7 days after RRT is discontinued.
  - ➢ patient has not died within 7 days after RRT is discontinued; post-RRT eGRF level not < 15 mL/minute/1.73 m^2 within 7 days.
- Partial AKI recovery after RRT, meets all of the following:
  - ➢ independence from RRT for > 72 hours (or 3 days when hours are not recorded)
  - ➢ last recorded SCr level > 1.5 times baseline level
  - ➢ return to a lower level of RIFLE criterion (eg, 3 to 2 or 1, or 2 to 1)

Appendix 2. Staging of acute kidney injury using modified RIFLE (KDIGO)13 recommendations

Stage	Serum creatinine criteria	Urine output criteria
1	Serum creatinine × 1.5 *or* serum creatinine rise of 0.3 mg/dL in 48 hrs	<0.5 mL/kg/hr for 6 hrs
2	Serum creatinine × 2	< 0.5 mL/kg/hr for 12 hrs
3	Serum creatinine × 3 *or* serum creatinine ≥ 4 mg/dL *or* had renal replacement therapy	< 0.3 mL/kg/hr for 24 hrs *or* anuria for 12 hrs

RIFLE = risk, injury, failure; loss, end-stage renal disease. KDIGO = Kidney Disease: Improving Global Outcomes.

Appendix 3. Trial sites, investigators and coordinators for the Protocolized Care for Early Septic Shock study

Advocate Christ Medical Center, Oak Lawn, Ill: E Kulstad, H Watts, K Hesse

Allegheny General Hospital, Pittsburgh, Pa: A Venkat, M Marynowski, S Livingston

Brigham and Women's Hospital, Boston, Mass: P Hou, A Massaro, S Parmar

Duke University Medical Center, Durham, NC: A Limkakeng, J Govert, W Drake

East Carolina University, Greenville, NC: T Delbridge, K Brewer, A Mainhart

George Washington University Medical Center, Washington, DC: A Dorfman, L Chawla, E Brasha-Mitchell

Intermountain Medical Center, Murray, Utah: C Grissom, T Allen, B Briggs

LAC+USC Medical Center, Los Angeles, Calif: H Belzberg, S Swadron, J Zhu

Louisiana State University Health Sciences Center, Shreveport, La: T Arnold, S Conrad, K Hutchinson

Maricopa Medical Center, Phoenix, Ariz: F LoVecchio, R Carlson, M Mulrow

Massachusetts General Hospital, Boston, Mass: M Filbin, A Waxman, B A Parry

Methodist Research Institute, Indianapolis, Ind: T Ellender, C Naum, C Lynn

North Shore University Hospital, Manhasset, NY: A Sama, T Slesinger, T Pastrana

Norwalk Hospital, Norwalk, Conn: J Fine, M Carius, C Belden

Penn State Hershey College of Medicine, Hershey, Pa: T Terndrup, M Wojnar, S Nafeei

Stanford University School of Medicine, Stanford, Calif: M Strehlow, R Pearl, V Ojha

Summa Health System, Akron, Ohio: S Wilber, B Martin, J Skruck

SUNY Downstate Medical Center, Brooklyn, NY: R Sinert, S Malhotra

Tampa General Hospital, Tampa, Fla: D Orban, R Paula, C Targal

Temple University Hospital, Philadelphia, Pa: J Ufberg, J Travaline, A Wang

UC Davis Medical Center, Sacramento, Calif: E Panacek, T Albertson, L Jones

University of Alabama at Birmingham, Birmingham, Ala: H Wang, K Lai

University of Arkansas for Medical Sciences, Little Rock, Ark: J Palmer, T Holmes, E Sides

University of Minnesota Medical Center, Fairview, Minn: N Schmiechen, C Weinert, S Nagamatsu

University of Pittsburgh Medical Center, Presbyterian Hospital, Pittsburgh, Pa: D Yealy, S Gunn, P Carey

University of Pittsburgh Medical Center, Shadyside Hospital, Pittsburgh, Pa: R Wadas, V Okwiya

University of Utah Health Sciences Center, Salt Lake City, Utah: E Kimball, E Harris, R Preston Vanderbilt

University Medical Center, Nashville, Tenn: W Self, D Dubinski

Washington Hospital Center, Washington, DC: M Goyal, C Phillips, R Migues

References

1. Rivers E, Nguyen B, Havstad S, et al. Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med. 2001; 345:1368–77. [PubMed: 11794169]
2. Reade MC, Huang DT, Bell D, et al. Variability in management of early severe sepsis. Emerg Med J. 2010; 27:110–5. [PubMed: 20156862]
3. ProCESS/ARISE/ProMISe Methodology Writing Committee. Harmonizing international trials of early goal-directed resuscitation for severe sepsis and septic shock: methodology of ProCESS, ARISE, and ProMISe. Intensive Care Med. Aug 20.2013 Epub ahead of print.
4. Bernard GR, Vincent JL, Laterre PF, et al. Efficacy and safety of recombinant human activated protein C for severe sepsis. N Engl J Med. 2001; 344:699–709. [PubMed: 11236773]
5. Abraham E, Reinhart K, Opal S, et al. Efficacy and safety of tifacogin (recombinant tissue factor pathway inhibitor) in severe sepsis: a randomized controlled trial. JAMA. 2003; 290:238–47. [PubMed: 12851279]
6. Warren BL, Eid A, Singer P, et al. Caring for the critically ill patient. High-dose antithrombin III in severe sepsis: a randomized controlled trial. JAMA. 2001; 286:1869–78. [PubMed: 11597289]
7. Knaus WA, Wagner DP, Draper EA, et al. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. Chest. 1991; 100:1619–36. [PubMed: 1959406]
8. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med. 1996; 22:707–10. [PubMed: 8844239]
9. Shapiro NI, Wolfe RE, Moore RB, et al. Mortality in Emergency Department Sepsis (MEDS) score: a prospectively derived and validated clinical prediction rule. Crit Care Med. 2003; 31:670–5. [PubMed: 12626967]
10. Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Med Care. 2005; 43:203–20. [PubMed: 15725977]
11. Jones AE, Shapiro NI, Trzeciak S, et al. Lactate clearance vs central venous oxygen saturation as goals of early sepsis therapy: a randomized clinical trial. JAMA. 2010; 303:739–46. [PubMed: 20179283]
12. Hosmer, DW.; Lemeshow, S. Applied logistic regression. 2nd ed.. John Wiley & Sons; New York: 2000.
13. Fitzmaurice, GM.; Laird, NM.; Ware, JH. Applied longitudinal analysis. John Wiley & Sons; Hoboken, NJ: 2004.
14. Cox DR. Regression models and life-tables. J Royal Stat Soc B. 1972; 34:187–220.

15. Gray RJ. Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. J Am Stat Assoc. 1992; 87:942–51.
16. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986; 73:13–22.
17. Hedeker D, Gibbons RD. Application of random-effects patternmixture models for missing data in longitudinal studies. Psychological Methods. 1997; 2:64–78.
18. Bekker PA. Comment on identification in the linear errors in variables model. Econometrica. 1986; 54:215–7.
19. Weinstein MC, Siegel JE, Gold MR, et al. Recommendations of the panel on cost-effectiveness in health and medicine. JAMA. 1996; 276:1253–8. [PubMed: 8849754]
20. American Thoracic Society. Understanding costs and cost-effectiveness in critical care: report from the second American Thoracic Society Workshop on outcomes research. Am J Respir Crit Care Med. 2002; 165:540–50. [PubMed: 11850349]
21. Angus DC, Linde-Zwirble WT, Clermont G, et al. PROWESS Investigators. Cost-effectiveness of drotrecogin alfa (activated) in the treatment of severe sepsis. Crit Care Med. 2003; 31:1–11. [PubMed: 12544986]
22. Angus DC, Clermont G, Watson RS, et al. Cost-effectiveness of inhaled nitric oxide in the treatment of neonatal respiratory failure in the United States. Pediatrics. 2003; 112(6 Pt 1):1351–60. [PubMed: 14654609]
23. McNamee R. Intention to treat, per protocol, as treated and instrumental variable estimators given non-compliance and effect heterogeneity. Stat Med. 2009; 28:2639–52. [PubMed: 19579227]
24. Bang H, Davis CE. On estimating treatment effects under noncompliance in randomized clinical trials: are intent-to-treat or instrumental variables analyses perfect solutions? Stat Med. 2007; 26:954–64. [PubMed: 16900559]
25. Lui KJ. Comments on ‘On estimating treatment effects under noncompliance in randomized clinical trials: are intent-to-treat or instrumental variables analyses perfect solutions?’. Stat Med. 2009; 28:531–4. [PubMed: 19125394]
26. Hewitt CE, Torgeson DJ, Miles JN. Is there another way to take account of noncompliance in randomized controlled trials? CMAJ. 2006; 175:347. [PubMed: 16908892]
27. Bonetti M, Gelber RD. Patterns of treatment effects in subsets of patients in clinical trials. Biostatistics. 2004; 5:465–81. [PubMed: 15208206]
28. Tibshirani R. Regression shrinkage and selection via the lasso. J Royal Stat Soc B. 1996; 58:267–88.
29. Bien J, Taylor J, Tibshirani R. A lasso for hierarchical interactions. Ann Statist. 2013; 41:1111–41.
30. Rubin DB. Multiple imputation after 18+ years. J Am Stat Assoc. 1996; 91:473–89.
31. Sattar A, Weissfeld LA, Molenberghs G. Analysis of non-ignorable missing and left-censored longitudinal data using weighted random effects tobit model. Stat Med. 2011; 30:3167–80. [PubMed: 21898524]
32. Lee M, Kong L, Weissfeld L. Multiple imputation for left-censored biomarker data based on Gibbs sampling method. Stat Med. 2012; 31:1838–48. [PubMed: 22359320]
33. Kidney Disease: Improving Global Outcomes (KDIGO) Acute Kidney Injury Work Group. KDIGO Clinical Practice Guideline for Acute Kidney Injury. Kidney Int. 2012; 2(Suppl):1–138.