Published in final edited form as:

Crit Care Med. 2013 July ; 41(7): 1674–1678. doi:10.1097/CCM.0b013e318287f850.

An Adaptive, Phase II, Dose-Finding Clinical Trial Design to Evaluate L-Carnitine in the Treatment of Septic Shock Based on Efficacy and Predictive Probability of Subsequent Phase III Success

Roger J. Lewis, MD, PhD^{1,2,3,4}, Kert Viele, PhD^{4,5}, Kristine Broglio, MS⁴, Scott M. Berry, PhD⁴, and Alan E. Jones, MD⁶

¹Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, CA

²Department of Emergency Medicine, Los Angeles Biomedical Research Institute, Torrance, CA

³Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA

⁴Berry Consultants, LLC, Austin, TX

⁵Department of Statistics, University of Kentucky, Lexington, KY

⁶Department of Emergency Medicine, University of Mississippi, Jackson, MS

Abstract

Objectives—Sepsis is the tenth leading cause of death in the United States. Despite extensive research, mortality rates for sepsis have not substantially improved in the last several decades. We describe an innovative phase II clinical trial design for evaluating the addition of L-carnitine to the treatment of vasopressor-dependent septic shock.

Design—The design incorporates a variety of features to increase efficiency, including a normal dynamic linear dose–response model, adaptive randomization, and early stopping for futility or success based on the probability that a future phase III trial using a 28-day mortality outcome would be successful.

Setting—Trial design and computer simulation of a trial to be conducted in the emergency department and ICU.

Interventions—Proposed to study intravenous L-carnitine.

Measurements—The proposed trial uses an early endpoint, the 48-hour change in Sequential Organ Failure Assessment score, to drive adaptive randomization and dose selection.

Main Results—We use existing data to model the expected relationship between the Sequential Organ Failure Assessment change and the 28-day mortality to determine the trial's operating characteristics using Monte Carlo simulation.

Conclusions—The resulting trial efficiently identifies the best dose of L-carnitine and provides clear guidance regarding whether to continue development into phase III.

Keywords

adaptive randomization; carnitine; clinical trial; sepsis

It is estimated that severe sepsis occurs with an incidence rate of three cases per 1,000 person-years, resulting in approximately 750,000 affected persons annually in the United States. Of those affected, approximately two thirds require ICU services (1). Sepsis ranks as the tenth leading cause of death in the United States with a hospital mortality rate of 30% resulting in 215,000 U.S. deaths annually (1, 2). During the last 30 years, several new therapeutic strategies for sepsis have been investigated. Unfortunately, these therapies have had little impact on the mortality rate in patients with sepsis, and the present in-hospital mortality rate approximately equals the mortality rate from the 1970s (3). These data underscore the need to investigate new therapeutic options for this devastating disease.

In a traditional clinical trial design, key characteristics are defined a priori and are not re-evaluated during the conduct of the trial. Examples of these trial characteristics include the assumed event rate in the control population and, in a phase II dose-finding trial, the region of the dose-response curve that is likely to be most clinically relevant. This approach may lead to substantial inefficiency or compromise important treatment goals. In a dose-finding trial with equal allocation across doses, for example, many of the subjects may ultimately be allocated to doses that are not in the clinically important region of the dose-response curve. These subjects yield little useful information about the doses of interest. These inefficiencies are particularly problematic in critically ill populations with high mortality rates, such as septic shock, because of ethical considerations, recruitment and consent challenges, and the time and resources required to perform trials in these populations.

In contrast, for a clinical trial that is adaptive by design, key trial characteristics evolve during the course of the trial according to predefined rules (4–6). One key trial characteristic that may evolve is the randomization probabilities used to allocate new subjects to the available treatment arms. For example, subjects can be preferentially allocated to the treatment arm that appears to be better performing, or similarly, subjects may be preferentially allocated to doses that appear increasingly to have clinically important benefits with an acceptable side-effect profile, as the emerging data identify those doses (5, 7–9). These adaptations improve both the efficiency of the clinical trial and the ethical balance between patient risk, benefit, and acquired information.

Although the use of surrogate outcomes is standard in phase II studies, it must be recognized that they are of less clinical importance compared with primary endpoints used in the phase III studies, such as 28-day mortality. Many previous, early-stage sepsis trials with positive findings using surrogate outcomes have failed to translate into clinically important improvements in patient-centered outcomes in later confirmatory phase III studies. However, the use of a surrogate outcome that is closely linked to the definitive outcome of interest and is rapidly available for each subject would allow preferential allocation of subjects to drug doses that are most effective while, simultaneously, allowing calculation of

the probability of success probability in a subsequent confirmatory trial based on the likely phase III endpoint. Using this methodology allows the optimal dose to be identified and only carried to phase III if the results are sufficiently promising.

Trial Design

Overall Design Strategy

We propose an adaptive, phase II trial to efficiently identify the dose of L-carnitine that provides the greatest improvement in the Sequential Organ Failure Assessment (SOFA) score (10) and simultaneously to assess the probability of success if that dose of L-carnitine was evaluated in a future phase III trial with a primary endpoint of 28-day mortality. To ensure we identify the best dose, a broad dose range will be considered (6, 12, and 18 g). If L-carnitine is effective, it is likely to exert its effects by improving multiple system organ function. Therefore, it is likely to have a direct effect on SOFA score and, accordingly, the change in SOFA score will be used to adaptively allocate subjects to the doses that are the most promising. This approach allows the initial consideration of a wide dose range but avoids the inefficiency of balanced randomization that would continue to allocate subjects to non-promising doses even late in the trial. Because our overall goal is to develop a treatment that decreases the mortality associated with vasopressor-dependent septic shock, the criteria for assessing the futility or success of this phase II trial will be based on the mortality benefit seen with the most-promising L-carnitine dose. Specifically, the trial will utilize frequent interim analyses, and based on the predictive probability of success in a future phase III trial, the trial will be stopped as soon as either: 1) a dose of L-carnitine has been identified that is sufficiently promising to warrant investigation in a phase III trial, or 2) none of the three doses are sufficiently effective to warrant further investigation.

Primary Outcomes, Response Adaptive Allocation, and Adaptive Design Plan

The trial will enroll a maximum of 250 subjects allocated to four treatment arms (control, 6, 12, 18 g). The adaptive trial will utilize both the change in SOFA score at 48-hour post-treatment and the 28-day mortality as efficacy endpoints. It should be noted that if an enrolled patient dies prior to calculation of the 48-hour SOFA score, a SOFA score at the time of death is calculated, and this value is used in lieu of the 48-hour value. The allocation of subjects will be determined only by the observed change in SOFA score; however, interpretation of the trial results will be based on both the change in SOFA score and the 28-day mortality. For determining the probability of success in a phase III trial (with 28-day mortality as the final endpoint), we assume the phase III trial would utilize standard, frequentist approaches and enroll up to a maximum of 2,000 subjects, with half in a control arm and half receiving the selected dose of L-carnitine. A phase III trial of that size would have a power of 95% to detect a reduction of 28-day mortality from 40% to 32% or a power of 78% to detect a reduction from 40% to 34% with L-carnitine therapy.

A normal dynamic linear dose-response model is utilized to improve the efficiency in estimating the dose response across dose levels for both of the efficacy endpoints (7, 11). During an initial “burn-in” period, 40 subjects are allocated equally among the treatment arms. From that point, an interim analysis is conducted after every 8 wk (approximately

every 12 enrolled subjects). The purpose of the interim analyses are two-fold, namely 1) to adjust the randomization proportions, so the likelihood of assignment to each active treatment arm is proportional to the probability that that arm leads to the greatest improvement in SOFA score at 48 hours; and 2) to evaluate whether the prespecified rules for early termination have been met. These analyses are a necessary part of adaptive trials, and the type I error rate and statistical power of the trial have been determined under the assumption that these steps are implemented. This interim analysis information will be transmitted to the data coordinating center to assure the randomization is updated for clinical sites; however, neither the sponsor nor the steering committee will have access to these interim results. However, the trial will be overseen by a traditional data and safety monitoring board (DSMB), working under the rules of an approved charter, and the DSMB will have access to unblinded interim enrollment, outcome, and safety data so that they are able to verify that the adaptive design is implemented and performing as designed and that the trial remains ethically and scientifically appropriate throughout its duration. Without access to unblinded data, effective oversight of the adaptive aspects of the trial would be impossible. After the burn-in period, a blocked randomization approach will be used to ensure that approximately one-third of subjects are allocated to the control arm throughout the trial. This helps to ensure that the result is resistant to confounding due to secular trends in outcome.

At each interim analysis after the 100th enrolled subject, the trial may also stop early for success or futility. Specifically, the trial is stopped for futility if there is less than a 40% posterior probability that the most promising L-carnitine dose leads to an improvement in SOFA at 48-hour posttreatment. The trial may stop early for success if there is a greater than 90% posterior probability that the most promising dose of L-carnitine improves SOFA at 48-hour posttreatment, and the predictive probability of demonstrating superiority in a future phase III trial of 28-day mortality is greater than 70%. The goal is to create strong success criteria that will stop the trial early only in situations where the best dose is well defined and success in phase III is highly likely. These criteria for defining futility and success were refined through trial simulations to yield desirable operating characteristics (see Trial Operating Characteristics section).

Trial Simulations

To estimate the likely relationship between the change in SOFA score at 48 hours and 28-day mortality, SOFA score and mortality data from more than 250 subjects with inclusion criteria identical to the proposed trial were obtained from a multi-center observational study. All subjects in the source study were treated according to the Surviving Sepsis Campaign guidelines for the management of septic shock (12). The observed mean change in SOFA score was 0.20 with a standard deviation of 2.69. The observed odds ratio for 28-day mortality, for each increase in SOFA of one point at 48 hrs, was 1.35 (95% confidence interval, 1.20–1.53). We used this relationship to generate virtual subjects for trial simulations, with L-carnitine having various hypothetical effects on the mean change in SOFA score and through this relationship on the 28-day mortality.

In order to evaluate the operating characteristics of the trial design, we conducted Monte Carlo simulations of the trial while assuming a variety of true dose–response relationships for both change in SOFA score and resulting 28-day mortality. For the purposes of simulation, subject enrollment was assumed to ramp up slowly over the first 12 weeks of the trial, reaching a plateau of 1.5 subjects per week. Under each of the scenarios listed in Table 1 (no, mild, and strong treatment effects), we simulated SOFA data for thousands of virtual subjects. Based on each simulated subjects' SOFA data, the background data was used to randomly assign a corresponding 28-day mortality outcome. From this large database of virtual subjects, we simulated thousands of possible trials. All trial design work and simulations were completed using the Fixed and Adaptive Clinical Trial Simulation Software from Berry Consultants, LLC, and Tessella (13).

Trial Operating Characteristics

Table 1 shows the resulting trial operating characteristics. Type I error is assessed by the probability of a positive trial under the null hypothesis that there is no relationship between L-carnitine dose and either SOFA score at 48-hour posttreatment or 28-day mortality. This trial has a type I error rate of 0.043. In this scenario, the trial stops early for futility with 43% probability, and the average sample size is 198.0 subjects. The trial was powered to detect an improvement in SOFA of two units at 48-hour posttreatment (see column entitled “Mild Treatment Effect” in Table 1). Under this scenario, the proposed trial has a power of 91.1% and stops early for success over two-thirds of the time. In this scenario, the average required sample size is 172.4 subjects, and the design is highly reliable in selecting the best dose to be carried forward. Moreover, it can be seen that under this scenario the trial efficiently allocates subjects to the best dose arm, maximizing the information gained per subject (Fig. 1). Table 1 also demonstrates the performance if there is a very strong treatment effect, in which case the power of the trial is even higher and the required sample size smaller.

Global Interpretation of Trial Results

At the end of the trial, the trial is considered negative if there is either less than a 90% posterior probability that the most promising dose of L-carnitine improves SOFA at 48-hour posttreatment or less than a 30% predictive probability that the most promising dose of L-carnitine would be successful in a subsequent phase III trial. The trial is defined as positive if there is a greater than 90% posterior probability that the selected dose of L-carnitine improves SOFA at 48-hour posttreatment and if there is a greater than 30% predictive probability of success in a future phase III trial. Considering the high mortality rate of septic shock and the potential benefit of a successful sepsis treatment, even a moderate probability of success in phase III may be worth pursuing.

Discussion

In this report, we describe an adaptive dose-finding clinical trial designed to evaluate L-carnitine in the treatment of vasopressor-dependent septic shock based on efficacy using a 48-hour SOFA score outcome and predictive probability of future phase III superiority using a 28-day mortality outcome. This allows the change in SOFA score to drive the response

adaptive allocation while 28-day mortality is investigated as the phase III endpoint. The use of this methodology allows the optimal dose of L-carnitine to be identified and only carried to phase III if the mortality results are sufficiently promising. In the proposed trial design, preliminary data on the association between a surrogate outcome (SOFA score) and the 28-day mortality were used to allow for realistic simulations of trial data. These adaptations improve both the efficiency of the clinical trial and the ethical balance between patient risk, benefit, and acquired information.

Designing a traditional phase II dose-finding clinical trial requires defining many key characteristics (e.g., the expected event rate in the control population, the region of the dose–response curve to be evaluated, and the randomization ratio to be used in allocating subjects to the treatment arms) prior to enrollment of the first subject. Accordingly, in a traditional design, the determination of operating characteristics such as type I error rate, power, and the required sample size is based on the assumption that these characteristics are held constant and that unplanned modification based on accumulating information is likely to increase the likelihood of a type I error. This traditional approach may lead to substantial inefficiency by allocating subjects to active experimental doses that are not in the clinically important region of the dose–response curve or by continuing to enroll subjects when, despite having some effect on a biomarker or surrogate endpoint, the agent has little chance of success in a confirmatory trial. Thus, these subjects could be considered in effect “wasted” as they yield little useful information about the doses of interest or have little effect on the chance of successfully completing a confirmatory trial. These inefficiencies also raise ethical considerations, particularly in populations with high mortality rates, and highlight other trial issues such as recruitment and consent challenges, and the time and resources required. The adaptive trial design proposed here directly addresses many of these challenges and improves the efficiency and the ethical balance of the trial.

Previous authors have recommended changes in the current paradigm of clinical trials conducted in the critically ill. Specifically, the use of surrogate outcomes for 28-day mortality and new clinical trial designs have come to the forefront of critical care research agenda (14, 15). In the design we propose, we utilize five key strengths of an adaptive trial design. First, a frequent interim analysis plan that allows ongoing assessment of evidence of treatment efficacy, rates of adverse events, and patient safety, all of which are important in this vulnerable population. Second, predefined decision rules determine whether the trial should be terminated early for success or futility. The third is response-adaptive randomization using a 48-hour SOFA score as the endpoint to inform the probabilities used to allocate subjects to the doses of L-carnitine. Fourth, a dose–response model is used to model the efficacy at the different doses allowing more efficient estimation. This method improves efficiency in locating the optimal dose of L-carnitine. Finally, we performed extensive trial simulations using historical data regarding the association between SOFA score and 28-day mortality to assure key performance characteristics (e.g., type I error rates, power, sample size, and accuracy in parameter estimation) are well characterized over a range of possible true treatment effects.

In our design, the trial result is considered a success if there is a greater than 90% posterior probability that the selected dose of L-carnitine improves SOFA at 48-hour posttreatment and

a greater than 30% predictive probability of success in a future phase III trial. Although this chance of success in phase III may be considered low, it was chosen for several reasons. First, septic shock is a disease with a considerably high mortality rate, and thus, successful sepsis therapies with even a moderate chance of success in phase III are worth pursuing. Second, L-carnitine has a very favorable safety profile and is inexpensive, thus making our first rationale even more compelling.

Conclusions

This report describes an adaptive phase II design for the investigation of adding L-carnitine to the treatment of vasopressor-dependent septic shock. The design incorporates a variety of innovative features to increase efficiency such as a normal dynamic linear dose–response model, responseadaptive randomization, and early stopping. The resulting trial efficiently identifies the best dose of L-carnitine while providing guidance concerning whether to continue development into phase III.

Acknowledgments

Dr. Lewis serves as senior medical scientist for Berry Consultants, serves as one of the principal investigators for a multiple PI UOI project supported by National Institutes of Neurological Disorders and Stroke, participates in EIR program of the U.S. FDA, CDHR, and serves as a consultant to AspenBio Pharma and Roche Diagnostics. He is also a board member of DSMBs for clinical trials sponsored by the National Institutes of Health (NIH), as well as Octapharma USA and Octapharma AG. Dr. Viele received funding from the NIH and is a consultant for Berry Consultants. Ms. Broglio is a consultant for Berry Consultants. Dr. Berry is owner of Berry Consultant, receiving money to consult on this trial execution. Dr. Jones's effort is supported by National Institute of General Medical Sciences/NIH (R01GM103799). He received grant support from Thermo Scientific.

References

1. Angus DC, Linde-Zwirble WT, Lidicker J, et al. Epidemiology of severe sepsis in the United States: Analysis of incidence, outcome, and associated costs of care. *Crit Care Med.* 2001; 29:1303–1310. [PubMed: 11445675]
2. Kochanek, K.; Smith, B. National Vital Statistics Report. Vol. 52. Centers for Disease Control and Prevention; 2004. Deaths: Preliminary Data for 2002; p. 1-32.
3. Centers for Disease Control and Prevention. Current trends increase in national hospital discharge survey rates for septicemia—United States, 1979–1987. *Morbidity and Mortality Weekly Report.* 1991; 39:31–34.
4. Gallo P, Chuang-Stein C, Dragalin V, et al. PhRMA Working Group. Adaptive designs in clinical drug development—An executive summary of the PhRMA Working Group. *J Biopharm Stat.* 2006; 16:275–283. discussion 285. [PubMed: 16724485]
5. Dragalin V. Adaptive designs: Terminology and classification. *Drug Information Journal.* 2006; 40:425–435.
6. Food and Drug Administration. [Accessed February 13, 2013] Guidance for industry: Adaptive design clinical trials for drugs and biologics. Available at: <http://www.fda.gov/downloads/Drugs/.../Guidances/ucm201790.pdf>
7. Berry SM, Spinelli W, Littman GS, et al. A Bayesian dose-finding trial with adaptive dose expansion to flexibly assess efficacy and safety of an investigational drug. *Clin Trials.* 2010; 7:121–135. [PubMed: 20338905]
8. Berry, DA.; Muller, P.; Grieve, AP., et al. Adaptive Bayesian designs for dose-ranging drug trials. In: Gatsonis, C.; Kass, R.; Carlin, B., et al., editors. *Case Studies in Bayesian Statistics V*. First. New York: Springer-Verlag; 2001. p. 99-181.
9. Bornkamp B, Bretz F, Dmitrienko A, et al. Innovative approaches for designing and analyzing adaptive dose-ranging trials. *J Biopharm Stat.* 2007; 17:965–995. [PubMed: 18027208]

10. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related organ failure assessment) score to describe organ dysfunction/failure. *Intensive Care Med.* 1996; 22:707–710. [PubMed: 8844239]
11. West, M.; Harrison, J. *Bayesian Forecasting and Dynamic Models*. Second. New York: Springer-Verlag; 1999.
12. Dellinger RP, Levy MM, Carlet JM, et al. International Surviving Sepsis Campaign Guidelines Committee; American Association of Critical-Care Nurses; American College of Chest Physicians; American College of Emergency Physicians; Canadian Critical Care Society; European Society of Clinical Microbiology and Infectious Diseases; European Society of Intensive Care Medicine; European Respiratory Society; International Sepsis Forum; Japanese Association for Acute Medicine; Japanese Society of Intensive Care Medicine; Society of Critical Care Medicine; Society of Hospital Medicine; Surgical Infection Society; World Federation of Societies of Intensive and Critical Care Medicine. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Crit Care Med.* 2008; 36:296–327. [PubMed: 18158437]
13. FACTS Development Team:FACTS. Austin, Texas: Berry Consultants and Tessella, Inc.; 2012. Fixed and Adaptive Clinical Trial Simulator. Available at: http://www.smarterclinicaltrials.com/wp-content/uploads/FACTS_introduction.pdf [Accessed February 13, 2013]
14. Angus DC, Mira JP, Vincent JL. Improving clinical trials in the critically ill. *Crit Care Med.* 2010; 38:527–532. [PubMed: 19851096]
15. Vincent JL. Endpoints in sepsis trials: more than just 28-day mortality? *Crit Care Med.* 2004; 32(5 Suppl):S209–S213. [PubMed: 15118519]

Figure 1. Results of simulations of 10,000 trials under the assumption of a mild treatment effect. As can be seen, subjects are preferentially allocated to the control and Dose 3 (18 g) arms.

Table 1
Operating Characteristics of Proposed Trial Design: Results of Monte Carlo Simulations (30,000 Simulated Trials)

Assumed treatment effects for simulations						
No Treatment Effect		Mild Treatment Effect		Strong Treatment Effect		
	SOFA	Mortality	SOFA	Mortality	SOFA	Mortality
Outcome: control	0	40%	0	40%	0	40%
Outcome: 6 g	0	40%	0	40%	-1	34%
Outcome: 12 g	0	40%	-1	34%	-2	28%
Outcome: 18 g	0	40%	-2	28%	-4	19%
Trial performance						
Probability of positive trial	0.043 (type I error rate)		0.911 (power)		0.999	
Probability of stopping early	For futility: 0.431		For futility: 0.001		For futility: 0.000	
	For success: 0.023		For success: 0.679		For success: 0.981	
Average required sample size	198.0		172.4		119.5	
Probability of selecting 18 g	0.35		0.99		1.00	
Average allocation of subjects between treatment Arms— <i>n</i> per arm (%)						
Control	62.7 (32%)		54.1 (31%)		36.5 (31%)	
6 g	47.0 (24%)		13.8 (8%)		10.5 (9%)	
12 g	38.7 (20%)		21.5 (12%)		12.5 (10%)	
18 g	49.6 (25%)		83.0 (48%)		60.0 (50%)	

SOFA = Sequential Organ Failure Assessment.