

Critical Care Perspective

Safeguarding Patients in Clinical Trials with High Mortality Rates

BRADLEY D. FREEMAN, ROBERT L. DANNER, STEVEN M. BANKS, and CHARLES NATANSON

Department of Surgery, Washington University School of Medicine, St. Louis, Missouri; and the Critical Care Medicine Department, National Institutes of Health, Bethesda, Maryland

As intensivists involved in both clinical and preclinical trials, we were struck by the outcomes of three recent investigations involving critically ill patients. These clinical trials, one examining human growth hormone in critical illness (1), the second, a blood substitute in acute hemorrhage (2), and the third, a nitric oxide synthase inhibitor in septic shock (3), showed with a high degree of significance (e.g., all with probability values (p) < 0.005) that the therapies under investigation were associated with increased rates of mortality (Table 1). Clinical investigations in critical illness are difficult undertakings due to patient heterogeneity, variations in practice standards, and high rates of organ dysfunction and mortality (4). Nevertheless, the demonstration that investigational therapy has caused a highly significant increase in mortality rates appropriately warrants a retrospective examination aimed at avoiding similar outcomes in future studies.

Our experience with performing drug trials in large animal models of septic shock sensitized us to the potential shortcomings of some clinical trial practices. Practices intended to expedite the conduct of clinical trials involving subacute diseases with low rates of mortality may prove unsafe when applied to acute, rapidly progressive, and highly lethal diseases (4). To safeguard patients with rapidly lethal diseases enrolled in clinical trials, we believe that toxicity monitoring in phase III trials should be unblinded and contemporaneous. Likewise, we believe that phase III studies should examine only drug dosing regimens that were determined to be safe in prior testing. However, in two of the cited studies, individuals blinded to treatment assignment attempted to perform toxicity monitoring using aggregate mortality data (1, 2). In one case, mortality rates were compared with historical controls (1). In the other study, a mortality rate predicted from a severity of illness-risk scoring system was used for comparison (2). Furthermore, in all three studies, phase I and/or phase II testing of the therapies as they were actually to be administered was not done (1–3).

Notably, none of these trial design practices was prohibited by current National Institutes of Health or Food and Drug Administration regulations or guidelines (5). Collectively, 150 more deaths occurred in the treatment arms than the control

arms of these three trials. As many as one-half of these deaths were in excess of those needed to show that the treatment was ineffective or harmful at accepted minimal thresholds for significance (e.g., $p < 0.05$). Equally important, neither the abstracts, published reports, nor accompanying editorials discussed aspects of study design or conduct that, if followed, may have prevented this outcome (1–3, 6, 7). Our purpose in writing this report is to examine factors of trial design and toxicity monitoring that appear to have contributed to an excessive number of treatment-related deaths. Further, we will suggest amendments to guidelines for conducting research in rapidly lethal diseases to help prevent such occurrences in future trials.

HUMAN GROWTH HORMONE IN CRITICAL ILLNESS

A recently published article in the *New England Journal of Medicine* studied the use of human growth hormone, an inhibitor of protein catabolism, in patients with critical illness. Two multicenter trials were simultaneously conducted—one based in Finland, and the other multinational, involving several European countries (1). The primary efficacy variable in these trials was intensive care unit length of stay. In addition, a number of secondary efficacy variables, including in-hospital mortality rates, were analyzed. At study completion, the authors demonstrated with a high degree of significance (e.g., $p < 0.001$ based on chi-square) that the administration of human growth hormone was associated with an increased mortality rate. When testing the null hypothesis that both the treatment and control groups had the same underlying mortality rate, Fisher's exact test (two-sided) yields a value of $p = 0.000004$ for the multinational study, $p = 0.001$ for the Finnish study, or based on the combined data from both studies, $p = 0.00000003$.

The authors of the phase III study did not perform pilot testing of human growth hormone in critically ill patients (1). Even in the absence of pilot testing, the finding that growth hormone administration significantly worsened survival rates should have been reached with much smaller studies and fewer deaths. This assumption is based on the following model: We examined a p value that was less significant than that reported for a harmful effect of human growth hormone (e.g., $p = 0.05$), and calculated the minimum number of patients needed to achieve this p value. We assumed that this p value was a threshold for toxicity, and therefore uncorrected. Further, we assumed a constant mortality rate, both during the 41 mo of patient accrual and among study sites. To estimate the number of excess patients enrolled, the number of patients required to reach a level of significance of $p = 0.05$ was subtracted from the number of patients entered in the particular study. Excess deaths were estimated by multiplying one-half the number of excess patients by the proportional increase in mortality rate that occurred in the treatment com-

(Received in original form November 9, 2000 and in revised form March 29, 2001)

Supported in part by GM00691-01.

The views expressed are the authors' and do not necessarily represent those of the National Institutes of Health, U.S. Public Health Service, or Department of Health and Human Services.

Correspondence and requests for reprints should be addressed to Bradley D. Freeman, M.D., Department of Surgery, Washington University School of Medicine, Suite 6104, Box 8109, St. Louis, MO 63110. E-mail: freemanb@msnotes.wustl.edu

Am J Respir Crit Care Med Vol 164, pp 190–192, 2001
Internet address: www.atsjournals.org

TABLE 1. MORTALITY RATES AND PROBABILITY VALUES FOR CLINICAL TRIALS IN CRITICAL ILLNESS

Therapy	Number of Patients	Mortality (%)		p Value	Reference
		Control	Treatment		
HGH*	532	19	42	< 0.001	1
DCLHb	112	17	46	< 0.003	2
L-NMA	797	49	59	< 0.005	3

Definition of abbreviations: DCLHb = diaspirin cross-linked hemoglobin; HGH = human growth hormone; L-NMA = L-N-methyl-L-arginine.

* Two separate multicenter trials were simultaneously conducted, one in Finland (n = 247) and the other multinational involving several European countries (n = 285).

pared with the control arm. To demonstrate a harmful effect for human growth hormone at the $p = 0.05$ level of significance, we estimate that the authors could have enrolled 154 fewer patients in the Finnish study, and 228 fewer patients in the multinational study. This potentially would have prevented a total of 44 deaths (e.g., 29 deaths in the multinational study and 15 deaths in the Finnish study). Even if we assumed that mortality rate was not constant, and that all treatment-related excess deaths occurred at the end of this 41-mo study, a level of significance of $p = 0.05$ for harm could have been achieved by enrolling 68 fewer patients, which would have prevented 8 deaths.

A close examination of this study suggests that the methods used to perform interim analysis and to monitor toxicity contributed to the occurrence of excess deaths (1). The initial plan to perform an interim analysis after enrolling 150 patients was not done due to slow accrual. Instead the investigators performed an analysis after enrolling 170 and 190 patients who could be evaluated in the Finnish and multinational study, respectively (1). The authors defined patients who could be evaluated as those who had “received growth hormone or placebo for three days and had survived for at least two days after discharge” (1). The published report states that “overall mortality rates and adverse events were continuously monitored” (1). The details of exactly how this was done were graciously provided by the first author of the published manuscript. The combined mortality rate of all patients in the study was compared with historical control subjects from a single participating institution by a safety officer from the sponsoring company who was blinded to treatment group assignment. These safety monitoring results were then internally reviewed by the company at 6-mo intervals and discussed with the investigators (J. Takala, personal communication). A Data Safety and Monitoring Committee was not used. Because control mortality rates were lower than predicted, when combined with the high mortality rates of the treatment arm, the average mortality rate was very similar to that of the historical cohort. The adverse effect of human growth hormone on survival rate was not appreciated until the planned fixed number of survivors had been enrolled and the study was unblinded and fully analyzed.

DIASPIRIN CROSS-LINKED HEMOGLOBIN (DCLHb) IN THE TREATMENT OF HEMORRHAGIC SHOCK

Problems related to pilot testing and monitoring of mortality rates also may have contributed to excess deaths in a phase III study involving patients with severe hemorrhagic shock (2). In this study, the blood substitute diaspirin cross-linked hemoglobin (DCLHb), a chemically modified human hemoglobin solution, was administered as adjunctive therapy. This study was designed to enroll approximately 850 trauma patients at 18 centers with presumed or proven hemorrhage and persistent hypoperfusion at the time of presentation. The primary

end point of this study was 28-d mortality rate. Interim analyses were planned after enrollment of 10%, 25%, 50%, and 75% of the accrual target, with the results of these analyses being subjected to prospectively defined stopping rules. This study was halted after 11 mo following the first interim analysis, at which time 112 patients had been enrolled, and 98 patients had received treatment or placebo. Twenty-eight day mortality rate in the 52 patients receiving DCLHb was 46%, compared with a mortality rate of 17% in the control arm ($p = 0.003$).

Like human growth hormone, DCLHb had undergone extensive prior examination. DCLHb had been evaluated in pre-clinical models of hemorrhagic shock (8), had undergone phase I testing in healthy volunteers (9), and had been administered to critically ill patients to evaluate its vasopressor effects (10). However, a pilot study of its use as an adjunctive therapy in patients with severe hemorrhagic shock, comparable with those enrolled in this phase III study, had not been done. Further, although a Data Safety and Monitoring committee was employed, this committee was blinded to treatment assignment and reviewed only “aggregate safety data” (2). The overall mortality rate in the patients receiving DCLHb or placebo (normal saline) was 33%, which did not significantly differ from the expected mortality rate of 40% (based on a risk prediction scoring system). Thus, as was the case with the studies of human growth hormone, the lethal effect of DCLHb was not detected until the treatment groups were unblinded, which in this case occurred at the time of the first interim data analysis. Using an analysis similar to that described above, we estimate that less than 42 patients could have been enrolled in this study, and that as many as 9 deaths related to DCLHb treatment could have been prevented, had a threshold for demonstrating a harmful therapeutic effect of $p = 0.05$ been used, and continuous unblinded or at least nonaggregated monitoring of toxicity taken place.

Notably, this was the first study implementing a Food and Drug Administration-approved mechanism (FDA regulation 21 CFR 50) in which investigators could be exempted from obtaining informed consent from the patients or their surrogates when it was not feasible to prospectively do so. Although it is always imperative to ensure the safety of subjects in clinical investigation, this issue achieves even greater importance when patients have not explicitly agreed to participate.

L-N^G-METHYLARGININE HCl (L-NMA) IN PATIENTS WITH SEPTIC SHOCK

L-NMA induces vasoconstriction through inhibition of the enzyme nitric oxide synthase. A prospective phase III study, involving 126 centers in both Europe and the United States, randomized patients with septic shock to receive either L-NMA or placebo in addition to conventional therapy (3). The primary efficacy variable in this study was 28-d survival, with secondary end points of 3- and 14-d survival. At an interim analysis of more than 500 patients, a trend toward a harmful effect for L-NMA was detected. However, by the time the interim analysis was completed and the decision had been made to halt the study, at least 250 additional patients had been enrolled. At study termination, a total of 797 patients had entered the trial, and a highly significant increase in mortality rates was demonstrated for L-NMA at all stated end points ($p < 0.005$) (3).

In contrast to preliminary evidence demonstrating the safety of human growth hormone and DCLHb in a variety of settings, preclinical testing of L-NMA in sepsis models had at least suggested the potential for serious toxicity (11–15). However, in a 32-patient open-label dose escalation study and in a larger phase II trial, L-NMA infusions for up to 7 d appeared

to have an acceptable safety profile in septic shock (16, 17). Based on the prospective plan of the subsequent phase III trial, the maximum duration of L-NMA infusion was doubled to 14 d in an effort to improve treatment efficacy while the study was underway (3). This maneuver effectively increased the total amount of drug delivered in the phase III trial without prior evaluation of this dose in humans.

L-NMA half-life is dose related and has been shown to increase during prolonged infusions (e.g., ≥ 8 h) (18). Thus, accumulation of L-NMA, a drug capable of producing myocardial depression and death at high doses in some preclinical sepsis models, may have accounted for its toxicity in this study. Testing a new drug regimen in critically ill patients for the first time in an international 126-center trial, as opposed to a more limited phase I or II study, is problematic. It is much more difficult to detect toxicity and rapidly halt enrollment before substantial patient accrual has occurred.

CONCLUSIONS

As demonstrated by these three studies (1–3), the number of treatment-related deaths occurring during the course of a clinical trial may be substantial. In an effort to reduce the number of such deaths in future studies, several additional guidelines appear necessary for conducting clinical trials in rapidly lethal diseases with high underlying mortality rates. All large phase III trials should be preceded by phase I or phase II studies of sufficient scope to estimate treatment effects, and to ensure that the therapy under investigation is not highly toxic. To accomplish these goals, pre-phase III studies need to be concurrently controlled and should administer the investigational drug in the same manner and to the same types of patients as those that will be included in the larger trial. Moreover, for safety monitoring purposes, comparisons between the treatment group and concurrent control group mortality rates should be in as close to real time as possible. Data Safety and Monitoring committees, or their independent surrogates, should never be limited to aggregate mortality rates and the use of historical controls. These measures, although followed in many investigations, are not uniformly required by Institutional Review Boards, nor mandated by the Food and Drug Administration or National Institutes of Health (5). Finally, if a trial demonstrates that a therapy is highly significantly associated with an increase in mortality rates, medical journals should require that the factors in the design and conduct of the clinical trial that led to this result be thoroughly discussed in the published report. Such a discussion is one of the best methods to disseminate important information, improve clinical trial design, and prevent future occurrences.

Acknowledgment: David Alling, M.D., Ph.D. (1918–1999) aided in the preparation of this manuscript through both his comments and statistical insights. Further, he was a friend and mentor to many of the authors.

References

1. Takala J, Ruokonen E, Webster NR, Nielsen MS, Zandstra DF, Vundelinckx G, Hinds CJ. Increased mortality associated with growth hormone treatment in critically ill adults. *N Engl J Med* 1999;341:785–792.
2. Sloan EP, Koenigsberg M, Gens D, Cipolle M, Runge J, Mallory MN, Rodman G Jr. Diaspirin cross-linked hemoglobin (DCLHb) in the treatment of severe traumatic hemorrhagic shock. *JAMA* 1999;282:1857–1864.
3. Grover R, Lopez A, Lorente J, Steingrub J, Bakker J, Willans S, McLuckie A, Takala J. Multi-center, randomized, placebo-controlled, double blind study of the nitric oxide synthase inhibitor 546C88: effect on survival in patients with septic shock. *Crit Care Med* 1999;27(1, Suppl): A33.
4. Palazzo M, Soni N. Critical-care studies: redefining the rules. *Lancet* 1999;352:1306–1307.
5. Code of Federal Regulations, Good Clinical Practice, Parts 50, 54, 56, 312, 314. Barnett International Press; 2000.
6. Bengissoon BA. Rethink about growth-hormone therapy for critically ill patients. *Lancet* 1999;354:2222–2223.
7. Demling R. Growth hormone therapy in critically ill patients. *N Engl J Med* 1999;341:837–838.
8. Gulati A, Sen AP. Dose-dependent effect of diaspirin cross-linked hemoglobin on regional blood circulation of severely hemorrhaged rats. *Shock* 1998;9:65–73.
9. Przybelski R, Kisicki J, Daily E, Bounds M, Mattia-Goldberg C, Colburn WA. Diaspirin cross-linked hemoglobin (DCLHB): phase I clinical safety assessment in normal healthy volunteers. *Crit Care Med* 1996; 24:1993–2000.
10. Reah G, Bodenham AR, Mallick A, Daily EK, Przybelski R. Initial evaluation of diaspirin cross-linked hemoglobin (DCLHb) as a vasopressor in critically ill patients. *Crit Care Med* 1997;25:1480–1488.
11. Cobb JP, Natanson C, Quezado ZMN, Hoffman WD, Koev CA, Banks S, Correa R, Levi R, Elin RJ, Hosseini JM, et al. Differential effects of L-NMMA in endotoxemic and normal dogs. *Am J Physiol* 1995;266: H1634–H1642.
12. Wright CE, Rees DD, Moncada S. Protective and pathological roles of nitric oxide in endotoxin shock. *Cardiovasc Res* 1992;26:48–57.
13. Laszlo F, Whittle BJ, Moncada S. Time-dependent enhancement or inhibition of endotoxin-induced vascular injury in rat intestine by nitric oxide synthase inhibitors. *Br J Pharmacol* 1994;111:1309–1315.
14. Nava E, Palmer RM, Moncada S. The role of nitric oxide in endotoxic shock: effects of NG-monomethyl-L-arginine. *J Cardiovasc Res* 1992; 20 (Suppl):S132–S134.
15. Nava E, Palmer RM, Moncada S. Inhibition of nitric oxide synthesis in septic shock: how much is beneficial? *Lancet* 1991;338:1555–1557.
16. Grover R, Zaccardelli D, Colice G, Guntupalli K, Watson D, Vincent JL. An open-label dose escalation study of the nitric oxide synthase inhibitor, N-methyl-L-arginine hydrochloride (546C88), in patients with septic shock. *Crit Care Med* 1999;27:913–922.
17. Grover R, Bakker J, McLuckie A, Anderson J, LeGall JR. The nitric oxide synthase inhibitor 546C88 promotes the resolution of shock in patients with severe sepsis. *Crit Care Med* 1998;26(1, Suppl):A29.
18. Hussein Z, Beerah M, Grover R, Jordan B, Jeffs R, Donaldson J, Zaccardelli D, Colice G, Guntupalli K, Watson D, Vincent JL. Pharmacokinetics of the nitric oxide synthase inhibitor L-N-G-methylarginine hydrochloride in patients with septic shock. *Clin Pharmacol Ther* 1999; 65:1–9.
19. DeMets DL, Pocock SJ, Julian DG. The agonising negative trend in monitoring of clinical trials. *Lancet* 1999;354:1983–1988.
20. Ponting GA, Halliday D, Teale JD, Sim AJW. Postoperative nitrogen balance with intravenous hyponutrition and growth hormone. *Lancet* 1988;1:438–440.
21. Ziegler TR, Rombeau JL, Young LS, Fong Y, Marano M, Lowry SF. Recombinant human growth hormone enhances the metabolic efficacy of parenteral nutrition: a double-blind, randomized controlled study. *J Clin Endocrinol Metab* 1999;74:865–873.
22. Knox J, Demling R, Wilmore D, Sarraf P, Santos A. Increased survival after major thermal injuries: the effect of growth hormone therapy in adults. *J Trauma* 1995;39:526–530.