# Clinical Trials in Septic Patients:

# Is a Single Trial Sufficient?

Internal Medicine Grand Rounds

Parkland Memorial Hospital

April 23, 1992

Robert S. Munford, M.D.

The future has never looked brighter for the management of septic patients. At least 8 potential therapeutics are undergoing evaluation at the present time (Table 1). Using a variety of strategies to interfere with the septic response cascade, these drugs bring hope that the risk of death from sepsis can be substantially reduced.

Table 1. Potential treatments for septic shock (1)

| Approach | Site of action | Company | Status (Oct. '91) |
|---|---|---|---|
| BPI | LPS | INCYTE | Preclinical |
| BPI | LPS | XOMA | Preclinical |
| Monoclonal Ab | LPS | Centocor | Awaiting FDA action |
| Monoclonal Ab | LPS | XOMA | Awaiting FDA action |
| Soluble IL-1ß | IL-1ß | Immunex | Preclinical |
| IL-1Ra | IL-1 receptor | Synergen | Phase II |
| Monoclonal Ab | TNFα | Cutter | Phase II |
| Monoclonal Ab | TNFα | Chiron, Bayer | Phase II/III |
| Soluble TNF receptor | TNFα | Immunex | Preliminary |
| Soluble TNF receptor | TNFα | Synergen | Preliminary |
| Monoclonal Ab | TNF-α receptor | Genentech, Roche | Preliminary |

BPI = bactericidal permeability-increasing protein. Source used by J. NIH Research: Mark Simon, Robertson Stephens & Co., San Francisco.

This Grand Rounds deals with a practical issue that faces clinicians as they deal with each of these agents: how to evaluate the results of drug efficacy trials in septic patients.

In this presentation I shall use seven of the major clinical trials of the past decade to illustrate key features of sepsis trials and to suggest some possible improvements. Some of the examples will be drawn from the analysis of the HA-1A monoclonal antibody trial presented at the Open Meeting of the Vaccines and Related Biological Products Advisory Committee, held at the Food and Drug Administration on September 4, 1991. The transcript of the meeting and the slides shown by Dr. Jay P. Siegel were obtained

through the Freedom of Information Office at the F.D.A. (HFI-35, F.D.A., 5600 Fishers Lane, Rockville, MD 20857. Telephone (301) 443-6310).

## The Sepsis Spectrum: Definitions

Sepsis is an exaggerated host response to microbial signal molecules. In many patients, sepsis follows a typical course, progressing through a recognizable spectrum of physiologic derangements from fever and tachycardia to shock and death. Since sepsis often occurs in individuals with major underlying diseases, however, its course may be unpredictably rapid or atypical. To improve communication among workers in the field, Bone (2) has proposed definitions for different stages of the sepsis spectrum:

*Bacteremia* is defined by the existence of positive blood cultures. Roughly 40 to 60% of patients with sepsis are bacteremic.

*Sepsis* is defined as "clinical evidence of infection, tachypnea (respiration > 20 breaths/min; if mechanically ventilated, > 10 L/min), tachycardia (heart rate > 90 beats/min), and hyperthermia or hypothermia (core or rectal temperature > 38.3°C or < 35.6°C)." In other words, "suspected infection plus the systemic response to it."

*Sepsis syndrome* is sepsis with evidence of altered organ perfusion, as evidenced by arterial hypoxemia, elevated lactate, or oliguria. Other workers might add altered mental status, or decreased SVR.

*Septic shock* is the sepsis syndrome with hypotension (systolic blood pressure < 90 mm Hg, or decrease from baseline systolic blood pressure > 40 mm Hg) that is responsive to fluids or inotropes.

*Refractory septic shock* is septic shock that "does not respond to conventional therapy within one hour."

In general, the risk of dying increases as one moves down the spectrum from sepsis to refractory septic shock. Although most studies have not clearly distinguished septic shock from refractory shock, "shock" clearly is associated with higher mortality than sepsis or sepsis syndrome (3-5)(and see below).

*Organ failure* in septic patients usually means ARDS, DIC, and/or renal failure. In one large study, these were more powerful predictors of death than was shock (6). This category, although not included by Bone (2), may also be useful for its prognostic value.

These definitions have not been uniformly praised; other workers have argued, for example, that the term "septicemia" should be retained to refer to patients who have

sepsis and bacteremia (as distinct from bacteremic patients who do not have sepsis) and patients with sepsis from local infections without bacteremia (7). The definitions are also somewhat imprecise (what is "conventional therapy"?).

## Clinical Trials in Septic Patients

Clinical trials in septic patients have been frustrated by two important kinds of heterogeneity. First, **patients who develop sepsis have many different predisposing factors** and, as discussed below, markedly different base line risks of dying. Supportive care and antimicrobial chemotherapy also differ between patients and study centers. It has been difficult to study drug efficacy in this diverse patient population. Second, **the etiologic agents that provoke the septic response are also heterogeneous**. Although gram-negative bacteria are isolated from most patients with sepsis, in 20 to 30% of septic patients the isolates are gram-positive bacteria or fungi. It is not possible to identify patients with gram-negative sepsis prior to treatment, so drugs that are specific for these bacteria (such as anti-endotoxin antibodies) will necessarily target a subgroup of the total population treated. Analyzing efficacy in population subgroups is hazardous, as will be discussed below. Drugs that interfere with the septic process itself--such as glucocorticoids, cytokine antagonists, and the like--would be expected to improve survival in the total population of patients treated, so that efficacy analysis may be much more straightforward for these drugs.

**The major clinical trials of the 1980's.**

**A. Antiendotoxin J5 antiserum in the treatment of Gram-negative bacteremia and shock (8).**

304 patients were randomized to receive either J5 antiserum or prevaccination serum from human volunteer vaccinees. Patients with gram-negative bacteremia were analyzed; patients with proved foci of gram-negative infection and negative blood cultures were included only if they had received appropriate antibiotic therapy prior to drawing blood cultures. J5 antiserum was associated with a significant reduction in mortality, even when patients were in shock at the time of infusion.

**B. Glucocorticoid therapy in patients with septic shock (9).**

59 patients with severe septic shock were randomized to receive high dose methylprednisolone, dexamethasone, or placebo. Patients who were treated with corticosteroids within four hours after the onset of shock had a higher incidence of shock reversal. There was no overall effect on mortality, however.

**C. High-dose methylprednisolone in the treatment of severe sepsis and septic shock (10).**

382 patients were enrolled from 19 centers to study the efficacy of high dose methylprednisolone (120 mg/kg) in patients with sepsis syndrome ± shock. No significant differences were found in the prevention or reversal of shock, or in the survival outcome. Significantly more deaths were related to secondary infection in the methylprednisolone group.

**D. Methylprednisolone in the treatment of early sepsis syndrome (11).**

223 patients were enrolled from 10 VA Hospitals. Patients with acute or chronic alteration of mental status were excluded because it was felt that informed consent should only be given by the patient. Although there was no effect of methylprednisolone (75 mg/kg) on survival in all patients with sepsis, there was a 75% reduction in mortality in the subgroup of patients with gram-negative bacteremia. No increased risk of side-effects was seen in the methylprednisolone group.

**E. Human IgG antibody to Escherichia coli J5 in the treatment of gram-negative septic shock (12).**

71 Swiss patients with septic shock were randomized to receive either a standard IgG preparation (IVIG) or a preparation of human IgG antibody to E. coli J5 (J5-IVIG). No difference between the groups was seen in mortality or the number of complications of sepsis.

**F. Human anti-endotoxin monoclonal antibody HA-1A in the treatment of gram-negative bacteremia (5).**

543 patients with sepsis syndrome ± shock were randomized to receive HA-1A, a human monoclonal IgM antibody to the lipid A region of endotoxin, or human albumin. Although there was no difference in overall mortality in the two groups, HA-1A was associated with lower mortality at 28 days in the subgroup of patients who had positive blood cultures for gram-negative bacteria. The drug was well tolerated and no anti-HA-1A antibodies were detected.

**G. Murine monoclonal anti-endotoxin antibody E5 in the treatment of gram-negative sepsis (4).**

486 patients with sepsis syndrome ± shock were randomized to receive E5, a murine IgM monoclonal antibody to the lipid A region of endotoxin, or human albumin. A survival advantage was found for patients with gram-negative sepsis who were not in shock at the time they were entered into the study. The drug seemed safe, although one-third or so of the patients who received E5 developed antibodies to murine IgM.

## The Analytical Plan

Investigators must now provide the F.D.A. with an Analytical Plan that will be used to analyze the data from the clinical trial of a drug. This plan typically has several important features:

**Study entry criteria.** An important advance during the past decade was the recognition that accurate assessment of sepsis therapy requires clinically precise entry and outcome definitions. The remarkably general definition used for enrolling patients in the landmark J5 antiserum trial ("patients were considered suitable...if they were severely ill, with recent deterioration in the form of sudden high fever or hypothermia, hypotension, or unexplained respiratory distress...)(8) has given way to much more quantitative entry criteria. These are summarized in Table 2.

Footnotes for Table 2:

♪ Refractory to infusion of at least 500 ml of 0.9% NaCl. "Decreased organ perfusion as evidenced by altered mental status or oliguria" and retrospective confirmation of "bacteremia or an identified source of infection" was required.

** At least four out of seven criteria within an eight-hour period, plus "clinical suspicion of sepsis." The seventh criterion was "a surgical or invasive procedure performed during the preceding 48 hours or the presence of an obvious primary septic site." A "normal sensorium" was required.

+ At least one of these four, plus "clinical evidence of infection" and all three of the other criteria shown.

♪ Presumed gram-negative septic shock, plus at least one of these signs (respiratory alkalosis was another).

* At least two of these four, plus one criterion for systemic septic response

@ Criteria for septic response in the E5 study; patients had to have one of these.

# In the presence of an "adequate fluid challenge" and the absence of antihypertensive agents. Entry required abnormal temperature, tachypnea, and tachycardia, plus either hypotension or two of six signs of systemic toxicity.

‡ Signs of systemic toxicity in the HA-1A study.

Case-fatality rates are for the placebo group in each study.

Table 2

**Comparison of entry criteria used in controlled clinical trials**

| Criterion | Sprung (9) | VA Cooperative (11) ** | Bone Methyl-prednisolone (10) | J5-IVIG (12) | E5 (4) | HA-1A (5) |
|---|---|---|---|---|---|---|
| Abnormal temperature | | yes | required | hypo-thermia♪ | yes* | required |
| Abnormal WBC | | yes | | | yes* | |
| Positive culture for gram-neg. | | | | yes | yes* | |
| "Gram-negative infection" | | | | yes | yes* | "suspected" |
| Tachycardia | | yes | required | | | required |
| Tachypnea | | yes | required | | yes@ | required |
| Hypotension | required♪ | yes | | required | yes@ | yes# |
| Metabolic acidosis | | | elevated lactate+ | yes♪ | yes@ | yes‡ |
| Decreased SVR | | | | | yes@ | yes‡ |
| Renal, CNS, or clotting dysfunction | oliguria<br>altered mentation | thrombo-cytopenia<br>normal sensorium | oliguria+<br>altered mentation+ | oliguria♪<br>clotting abn.♪ | yes@ | yes (each given equal weight)‡ |
| Arterial hypoxemia | | | yes+ | yes♪ | | yes‡ |
| Case-fatality rate at 14 days (%) | 69 | 21 | 25 | 54 | 28 | 43 (28 days) |

Each of the trials obviously used different entry criteria. In effect, **the entry criteria select patients with a certain spectrum of disease severity. The drug will only be tested in these patients.**

For example, in the VA Cooperative Sepsis Study, it was felt that only the septic patient could give informed consent to participate in the study. As a result, approximately 75% of the patients who were screened for the study had to be excluded. This entry criterion selected patients who had less severe sepsis, as indicated by the relatively low case fatality rate in the placebo group (21%). A retrospective analysis of the VA data indicated that patients who had acutely altered mental status had a case fatality rate of 49%, twice that of the patients with normal mental status (13). Patients with acutely altered mental status were also significantly more likely to have hypotension or thrombocytopenia, two parameters that have been associated with risk of death from sepsis. A drug that is shown to be efficacious using these selection criteria may not work in patients with more severe sepsis.

Other entry criteria may also exclude certain septic patients. For example, in the same VA study, 26% of 1,623 screened patients who met the study definition of sepsis had normal temperatures (between 96°F and 101°F)(7) and would not have been enrolled in clinical trials, such as the HA-1A (5) and High Dose Methylprednisolone (10) studies, that required an altered temperature for entry. In the VA study, patients with normal temperature had a higher case fatality rate (44%) than those with fever (28%).

Studies that target patients with septic or refractory septic shock will obviously exclude patients with earlier stages of sepsis.

So study entry criteria may select, from the total population of septic patients, subjects with early or late septic responses. As shown in Table 2, different selection criteria are associated with different risks of dying from the septic episode. The efficacy of the study drug will only be tested in patients who meet these selected criteria. If a drug is tested in a carefully defined subset of septic patients, it may only work in such patients. Using a sepsis drug in patients who do not meet the entry criteria used in the pivotal clinical trial of that drug may be experimental therapy, and physicians should not be coaxed by zealous manufacturers to "extend" favorable results in one group to other groups in which the efficacy of the drug has not been shown.

Although each set of treatment criteria could be used by clinicians in practice, the **non-uniformity in clinical definitions and study entry criteria makes it very difficult to compare the results of different studies.** If two studies define shock differently, how is one to compare a drug that evidently only works in patients who are in shock (in study A) with another drug that only works in patients who are not in shock (in study B)? Such is the situation with the E5 and HA-1A studies. The HA-1A trial defined shock as "systolic blood pressure of less than 90 mm Hg or the use of vasopressor drugs to maintain blood

pressure,"(5) while the E5 study used a definition of shock ("refractory" shock) that required "organ dysfunction with systemic hypotension (systolic blood pressure < 90 mm Hg) refractory to fluid resuscitation (minimum, 500 mL of isotonic fluid) and inotrope therapy if used." Many patients with "shock" according to the HA-1A definition would probably be included in the "no shock" subgroup of the E5 study. A comparison of the results of these studies is thus limited to simple, non-quantitative generalizations.

**Study outcome criteria.** End-points must also be designated before the study is performed or analyzed. Almost all of the recent studies have used the most relevant and ascertainable end-point, death. There are two important issues, however. First, which deaths should be counted? Second, how long should the patients be observed?

**Septic and all-cause mortality.** "Septic" deaths include deaths that are deemed the result of organ failure produced by the septic episode. In some studies (8,14), but not others (4,9,11), septic and non-septic deaths were considered separately. In the J5 antiserum trial (8), for example, deaths that were deemed to be "from a cause entirely unrelated to bacteremia," i.e., "if they took place several weeks after shock and infection had resolved," were not considered in the efficacy analysis. In contrast, the HA-1A study protocol provided that both septic and all-cause mortality would be considered. (Deaths were attributed to sepsis or other causes prior to breaking the code.) Whereas HA-1A was associated with a statistically significant reduction in all-cause mortality at 28 days, the difference in septic mortality was not significant. As it happened, four patients in the placebo group died of non-septic causes during the last two weeks of the observation period, providing the difference necessary for statistical significance in the "all-cause" mortality group. So the drug did not definitely prevent death from the disease it was intended to treat, yet it may have had a beneficial effect on outcome when all-cause mortality was counted. Everyone would agree that a sepsis therapeutic should prevent deaths attributable to sepsis, whereas the ability of such drugs to prevent non-septic deaths is more problematic. On the other hand, all-cause mortality is easier to determine.

**The length of the observation period.** When should the outcome of a sepsis therapeutic be evaluated? Most studies used a 14 day end-point, since most septic deaths were thought to occur before this time. In both the E5 and HA-1A studies, a longer (28 or 30 day) end-point was also used. Interestingly, mortality increased substantially between 14 and 28/30 days in both of these studies:

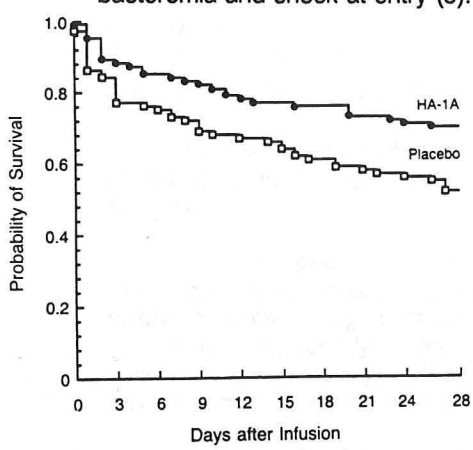
Table 4.

**Placebo group mortality at 14 and 28 days**

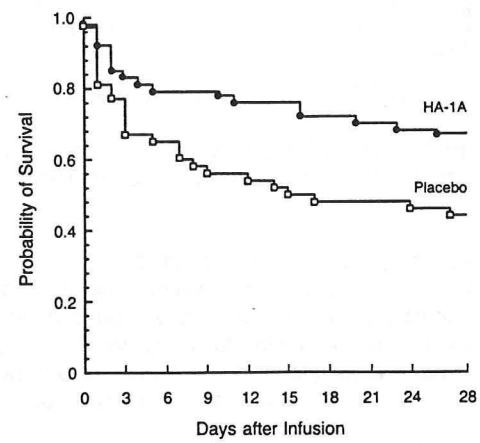
| Clinical trial | 14 day mortality (%) | 28-30 day mortality (%) |
|---|---|---|
| **E5** (all patients) | 28 | 41 |
| **HA-1A** (gram-negative bacteremia subgroup, all cause mortality) | 34 | 49 |

These data suggest that observation for 28-30 days may give a better estimate of the impact of sepsis. The longer the period of observation, the more likely one may be to detect mortality due to sepsis-induced organ failure. Definitive attribution of a death to the septic episode may become more difficult as time passes, however. It is difficult to know whether "non-septic" deaths that occur during the 14 to 28 day interval should be counted for (or against) a study drug, although randomization would generally be expected to reduce the risk of bias.

**Time trends.** Survival data are commonly shown graphically as Kaplan-Meier plots. This format gives the reader an appreciation for the actual distance between the placebo and study drug curves over time. Other methods for determining significant differences between study groups may be somewhat misleading. For example, in published account of the HA-1A trial, the survival curves were shown for the total subgroup of patients with gram-negative bacteremia and for patients with gram-negative bacteremia and shock at entry (5).




Survival of patients with gram-negative bacteremia




Survival of patients with gram-negative bacteremia and shock

Regarding the patients with gram-negative bacteremia who were not in shock, only the data from day 28 were described. It was stated that a "Cox proportional-hazards model was fitted to the survival data with treatment and shock as independent variables. This analysis indicated that shock was an important determinant of survival (P = 0.047) and that HA-1A reduced mortality in both patients with shock and patients without shock (P = 0.012)"(5). In contrast, the F.D.A., analyzing the time trend in the same data, came to the following conclusion: "in the no-shock group there was not a significant difference in mortality...If you look over time, those curves cross. At some time periods it is higher in the treatment group and, in others, in the placebo. There does not appear to be a significant difference"(15).

The survival data at 14 and 28 days after infusion are shown in the following Table:

Table 5.

**Effect of shock on the efficacy of HA-1A at 14 and 28 days after infusion**

| | 14 day mortality** | 28 day mortality |
|---|---|---|
| **SHOCK** | | |
| Placebo | 23/48 (48%) | 27/47 (57%) |
| HA-1A | 13/54 (24%) p = 0.012 | 18/54 (33%) p = 0.017 |
| **NO SHOCK** | | |
| Placebo | 9/47 (19%) | 18/47 (38%) |
| HA-1A | 12/51 (24%) p = 0.6 | 14/51 (27%) |

** dead/total (%). Day 14 data from slide presentation by Dr. Jay P. Siegel (14)

This experience makes a strong case for requiring analysis at intermediate time points (using, for example, the Kaplan-Meier format) in published reports of trials in septic patients.

**Soft end points.** Some studies have also reported soft end points, such as resolution of morbidities, time to hospital discharge, and the like. These end points often involve subjective judgments and can be taken seriously only when they are clearly described prospectively in the analysis plan. They are probably more useful as secondary (confirmatory) end points than as the primary end point--sepsis is a lethal disease, and the primary goal of therapy should be to prevent death.

**Safety end points** are also part of the Analytical Plan. The monoclonal antibody trials were concerned that the patients might make antibodies to the infused materials and with possible idiosyndratic (anaphylactic) reactions to the proteins. Concern was

expressed that HA-1A may have been associated with increased mortality in patients without gram-negative bacteremia (16), since in this group there was slightly higher mortality (43%) in the HA-1A group than in the placebo group (39%). This finding, while not statistically significant and probably of little consequence, raises the point that it may be difficult to detect a 1-2% incidence of lethal toxicity in a drug that is used to treat patients whose baseline risk is so heterogeneous (see below).

**Efficacy Subgroups**

No component of the Analytical Plan is more critical than the designation of the Efficacy Subgroups that will be analyzed. Put simply, the more subgroups studied, the more chances to show efficacy and the more stringent the criteria for statistical significance must be.

The VA Systemic Sepsis Cooperative Study Group described its subgroup analysis in the published report (11). There were three patient subgroups: those with evidence of sepsis (205 patients), those with gram-negative bacteremia (51 patients), and those with gram-negative infections (136 patients). All-cause mortality was determined at 14 days. P values (two-sided) less than 0.01 were considered to be significant. A sequential monitoring procedure was based on the detection of a two-thirds reduction in mortality with glucocorticoid therapy, with a Type I error rate of 5 per cent and a Type II error rate of 5 per cent overall and 20% in the gram-negative bacteremia group.

The 14 day mortality in the three subgroups was as follows:

Table 6.

**The VA Systemic Sepsis Cooperative Study of Glucocorticoid Therapy: 14 day all-cause mortality (11)**

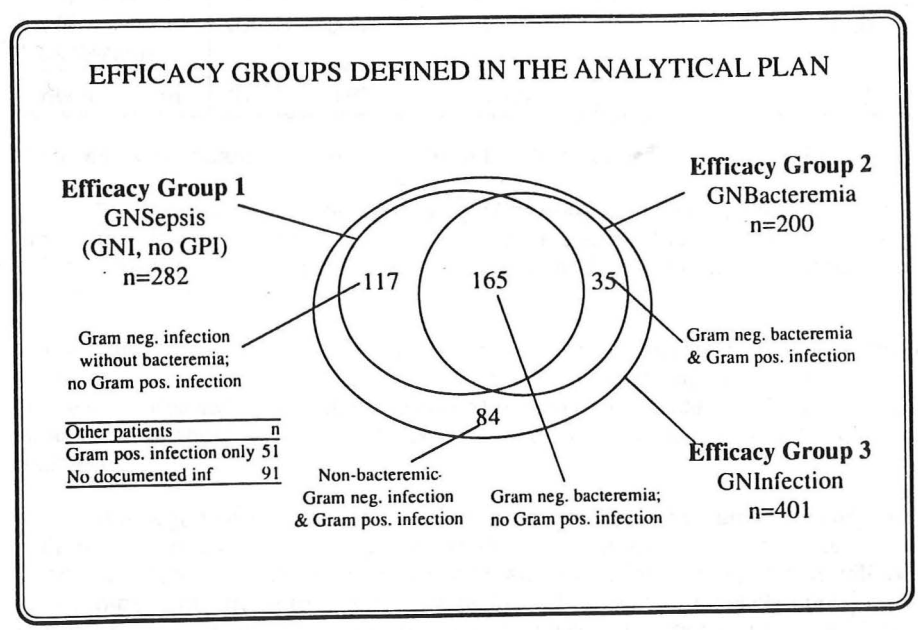
| Subgroup | n | Placebo | Glucocorticoid | P value |
|---|---|---|---|---|
| **Evidence of sepsis** | 205 | 21/98 (21.4)** | 20/107 (18.7) | 0.75 |
| **Gram-negative bacteremia** | 51 | 6/22 (27.3) | 2/29 (6.9) | 0.11 |
| **Gram-negative infection** | 136 | 17/67 (25.4) | 12/69 (17.4) | 0.35 |

** dead/total (%)

No significant differences were found, although a tendency in favor of glucocorticoid therapy was found in the gram-negative bacteremia group.

According to the FDA analysis, the HA-1A analytical plan provided for 3 Efficacy

Groups, each of which was analyzed according to two kinds of mortality ("Septic" and "All-Cause") at two times following infusion (14 and 28 days), resulting in 12 subgroups. The Efficacy Groups were Gram-negative sepsis (documented gram-negative disease, no infection with other microbes), gram-negative bacteremia (positive blood culture for gram-negative bacteria, with or without positive cultures for other microbes), and gram-negative infection (all patients with gram-negative disease, regardless of other kinds of ongoing infections). The distribution of patients in these categories is shown in the following diagram from a slide shown by Dr. Jay P. Siegel at the F.D.A. open meeting of the Vaccines and Related Biological Products Advisory Committee, F.D.A., September 4, 1991 (14).



It was determined by the F.D.A. statisticians that, because the subgroups were somewhat overlapping, the criterion for statistical significance should be somewhere between 0.01 and 0.03 (rather than 0.05)(15). Only one of the subgroups met this criterion: Gram-negative bacteremia, all-cause mortality, 28 days (Table 7). This subgroup was the focus of the published account of the trial (5).

Table 7.

**The HA-1A trial: FDA analysis of treatment effect on mortality**

| | | at day 14 | | over 28 days | |
|---|---|---|---|---|---|
| | n | septic mortality | all-cause mortality | septic mortality | all-cause mortality |
| **GN sepsis** | 282 | 0.56 | 0.56 | 0.29 | 0.18 |
| **GN bacteremia** | 200 | 0.12 | 0.12 | 0.039 | 0.014 |
| **GN infection** | 401 | 0.89 | 0.89 | 0.47 | 0.30 |

P values - Chi-square analysis. Data from F.D.A. analysis by Dr. Jay P. Siegel.

The number and nature of the subgroups analyzed in the E5 study is less certain, since the subgroup analysis was not described in detail in the available reports and the F.D.A. analysis of the E5 data was not presented (except "for discussion") at the September 4, 1991, meeting.

**Covariates.** Another important feature of the Analytical Plan is the choice of Covariates. These are possible confounding variables that might influence the outcome of the trial, or clinical parameters that could be used to administer the drug--the presence or absence of shock or altered mental status at the time of infusion are the most obvious examples.

In a large trial, randomization should produce groups that are essentially identical with regard to various baseline and treatment parameters. In practice, however, **randomization does not always achieve an equal distribution of patients with regard to all important characteristics.** The problem becomes particularly important when subgroups are carved out of the larger population--a random distribution of patients according to some baseline characteristic may be present in the larger group, but not in smaller subsets. This is a problem that has plagued sepsis trials and deserves particular attention here. There are four difficult areas: the **severity of underlying illness**, the presence or absence of **shock**, the **adequacy of antimicrobial therapy**, and the **duration of sepsis prior to treatment**.

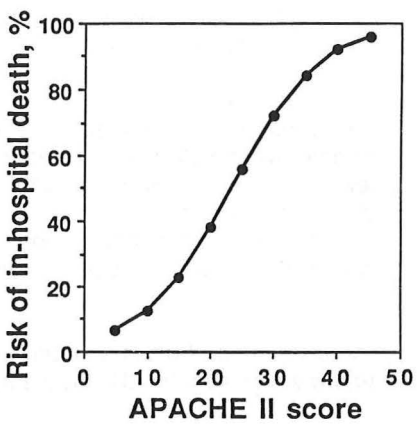
**Severity of Underlying Illness.** The relationship between the risk of death from gram-negative bacteremia and the severity of a patient's underlying illness was noticed by McCabe and Jackson (17) and used for many years in studies of sepsis. Recently, this simple categorization scheme has been superceded by multivariable prognostic scoring systems. The APACHE (Acute Physiology, Age, Chronic Health Evaluation) method evaluates patients according to three parameters (age, comorbid conditions, and physiologic variables) at the time of admission to the ICU and combines these with weighted variables (disease category) and selection criteria (location prior to ICU admission) to derive a prognostic score (18,19). Drawing on a database of 17,440 adults admitted to medical and surgical ICUs, the updated APACHE III scoring system provides a predicted risk of in-hospital death for individual patients (19).

It is important to recognize that **the relationship between APACHE score and risk of hospital death is non-linear** (18,20)(see Figure below), and that **the slope of the curve is steepest over the range of APACHE scores found in septic patients.** Moreover, the equation used to determine the risk of hospital death incorporates, in addition to the APACHE score, variables such as emergency surgery, location in the hospital prior to transfer to the ICU (patients transferred from the emergency room fare better than those transferred from wards or other ICUs, etc.), and disease category. Simply comparing mean APACHE scores in the placebo and treatment groups, as was done in the E5 and HA-1A studies, gives an incomplete and possibly inaccurate appreciation of their true prognostic stratification. The individual risks of death should be calculated and used instead of the raw APACHE scores.

Risk of hospital death (R):

$$\ln(R/1-R) = -3.517 + (\text{APACHE II score} \times 0.146) + (0.603 \text{ if post-emergency surgery}) + (\text{Diagnostic category weight})$$

For sepsis, the diagnostic category weight is 0.113. See references 18 and 19.



For example, the APACHE II scores in the HA-1A study were 25.7 ± 8.1 (placebo) and 23.6 ± 9.0 (HA-1A)(5). The risk of hospital death, calculated using the above equation and assuming that there was no emergency surgery, was 58.7% ± 28% (placebo) and 51.9% ± 30% (HA-1A). Citing only mean APACHE II scores obscures the base line physiological heterogeneity within the study groups and minimizes their apparent differences. The problem is particularly important when the APACHE score is in the middle range, where the risk of death curve is steepest.

The APACHE system has been criticized because it does not include clotting tests among the physiological variables used to tabulate the score (DIC, an identified risk factor for death in septic patients, is thus not included in the APACHE score). The risk of death equation incorporates a constant factor for patients with sepsis and it is also not certain that the APACHE system can accurately predict the different risks of death for patients with septic shock vs. those with refractory septic shock. In addition, the APACHE system obviously does not take into account important post-stratification variables such as the adequacy of supportive care or antimicrobial chemotherapy.

In addition to comparing multivariable prognostic scores in the treatment and placebo groups, one would like to know that the groups were reasonably similar with regard to important risk factors for death in septic patients. Here it is commonly stated that the "differences between the study groups were not significant." In fact, statistically significant differences between the groups are not really the point--the issue is whether or not the groups were imbalanced in such a way that the observed treatment effect might be spurious. As Dr. Robert Haley points out, **"it doesn't matter whether the imbalance occurred by chance, it matters whether it produced a bias."** Pocock et al. (21) reached a similar conclusion: "base-line differences should not be detected by significance testing, since the effect of a prognostic factor on the overall difference in treatment results depends both on its effect on response and the magnitude of the imbalance between groups."

In the HA-1A study, the placebo and HA-1A groups were well matched for hypotension (51% in each group), yet the placebo group had a higher prevalence of ARDS (13% vs. 9%), acute hepatic failure (26% vs. 19%), acute renal failure (46% vs. 35%), and DIC (21% vs. 18%)(5). Since the overall treatment effect found in this study was borderline, the imbalances in these base-line factors, several of which are thought to increase the risk of death from sepsis, cast serious doubt on the validity of the advertised trial result.

**Shock.** As discussed above and documented in numerous studies, the existence of hypotension greatly increases the risk of dying from sepsis. Unfortunately, **the major clinical trials defined shock in different ways.**

Table 3.

**Definitions of shock used in clinical trials**

| **Septic shock** | |
|---|---|
| J5 | not stated. "Profound shock" was defined as requiring pressors for more than 6 hours |
| VA Cooperative | supine position, cuff pressure in the upper arm < 90 mm Hg systolic |
| Bone | sustained decrease in systolic blood pressure to less than 90 mm Hg, or a drop of 40 mm Hg from base line, for at least one hour; with adequate volume replacement and no antihypertensive medication |
| HA-1A | systolic blood pressure of less than 90 mm Hg or the use of vasopressor drugs to maintain blood pressure |
| J5IG | systolic blood pressure <90 mm Hg or a decrease of >30 mm Hg in a hypertensive patient in the absence of other causes of shock |
| **Refractory septic shock** | |
| Sprung | systolic blood pressure under 90 mm Hg or 50 mm Hg less than a previously defined pressure in a hypertensive patient, continuing despite an iv infusion of at least 500 ml of 0.9% sodium chloride. |
| E5 | organ dysfunction with systemic hypotension (systolic blood pressure < 90 mm Hg) refractory to fluid resuscitation (minimum, 500 mL of isotonic fluid) and inotrope therapy if used |

Most studies required a systolic blood pressure $\leq$ 90 mm Hg, or a substantial fall in systolic pressure in previously hypertensive individuals. Beyond this basic definition, the studies differed according to whether the definition of shock required (a) sustained hypotension, lasting some predetermined time period, (b) hypotension that was refractory to a fluid challenge (and how much fluid is required before the definition was met), and (c) whether pressor support was necessary (and if so, how much). In some, evidence of organ hypoperfusion was also required. Since the outcome of reversible and refractory shock differs substantially, the definition of shock used to categorize patients in a particular study can be very important. (Further description of the severity of shock, using dopamine requirement and/or blood lactate levels, may improve the prognostic precision of the refractory shock definition: a small study that enrolled only patients with refractory

shock found that 11 of 16 (69%) patients in the placebo group died (9). When the severity of shock was assessed using lactate levels and dopamine requirements, patients with severe shock had higher mortality rate (91%) than patients with moderate (67%) or mild (50%) shock (9).)

**Adequate Antimicrobial Therapy.** There is strong evidence that appropriate antimicrobial therapy benefits patients with sepsis. A retrospective 10-year review from the Boston City Hospital, published a decade ago (3), found that appropriate antimicrobial therapy reduced the risk of shock in patients with gram-negative bacteremia regardless of the patient's underlying disease:

TABLE 8.

**Relationship of antibiotic therapy to development of shock (3)**

| Underlying Host Disease | Antibiotic Therapy at Onset of Bacteremia | |
|---|---|---|
| | Appropriate | Inappropriate |
| Rapidly fatal | 7/25 (28%)[+] | 6/9 (67%) |
| Ultimately fatal | 33/140 (24%) | 53/103 (51%) |
| Nonfatal | 28/154 (18%) | 29/67 (43%) |

[+] number with shock/number of patients with bacteremia (% with shock)

Overall, in this series the case-fatality rate was 41% when patients with septic shock received appropriate antimicrobial therapy (i.e., at least one effective drug was given) and 59% when antimicrobial therapy was inappropriate.

Other studies have reached similar conclusions (22). In the placebo arm of the HA-1A study, patients with gram-negative bacteremia who received appropriate antimicrobial therapy within 24 hours of study entry had a case-fatality rate of 27% (21 of 79); 69% (11 of 16) of those who received inappropriate antimicrobial therapy died (15). In this study, adequate antimicrobial therapy was defined as the administration of "an antibiotic to which each isolated organism was sensitive" within one day before or after infusion of the study material. It is not known whether the patients who received adequate and inadequate antimicrobial therapy were well matched for other important parameters such as underlying disease.

One recent study did not find that adequate antimicrobial therapy improves survival in septic patients (23); this study, performed in a community hospital, may have involved patients who were less severely ill than those in the various studies from university hospital centers. Another study, while confirming an important role for antimicrobial therapy, found that the therapy administered on the first day blood cultures were positive

was not as important as therapy given subsequently (24).

The importance of adequate antimicrobial therapy as a covariate in sepsis therapy trials increases when the study agent requires bacterial lysis for its efficacy. For example, the HA-1A monoclonal antibody is now said to bind the LPS of most smooth bacteria (i.e., most clinical isolates) only when the bacteria have been treated first with an antimicrobial (15,25). If this is true, it is unlikely that this antibody would work in patients who do not receive appropriate antimicrobial therapy--so such patients should probably be excluded from the analysis altogether. When one does this, unfortunately the difference between the placebo and HA-1A groups at 14 days after infusion decreases substantially:

TABLE 9.

**Adequacy of empiric antibiotic therapy in the HA-1A trial**
**Subgroup analysis in patients with gram-negative bacteremia**
**- mortality at 14 days[**]**

| | Placebo | HA-1A | p value |
|---|---|---|---|
| **All patients** | 32/95 (34%) | 25/105 (24%) | 0.12 |
| **Inadequate antibiotics** | 11/16 (69%) | 5/10 (50%) | > 0.4 |
| **Adequate antibiotics** | 21/79 (27%) | 20/95 (21%) | > 0.4 |

** deaths/total (%)

(Adapted from data presented by Dr. Jay P. Siegel, the F.D.A. analyst (14).)

This and other covariates may also be considered using multivariate analysis. When this approach was used for the HA-1A data, adjustment for adequate antimicrobial therapy raised the p value by 2- to 3-fold in each category. The p value for all-cause mortality at 28 days in patients with gram-negative bacteremia increased from 0.014 to 0.033 (14).

In the E5 study, "antibiotic therapy was considered appropriate when all isolates tested were sensitive to at least one antibiotic administered" during the first 3 days after entry (4). 94% of E5-treated patients and 99% of placebo patients were said to have received appropriate antimicrobial therapy. The VA Cooperative study (11) classified patients with positive cultures upon entry as receiving "appropriate" antibiotic therapy "if sensitivity studies indicated that the organism or organisms were susceptible to one or another of the initial drugs used; antibiotic therapy was considered "inappropriate" when the organism or organisms were not sensitive to any drug used in the empirical antibiotic

regimen." Ninety-one percent of the placebo recipients and 83% of the glucocorticoid recipients received appropriate antimicrobial therapy. There was no difference in outcome between those who did and did not receive appropriate antimicrobials in this study of early sepsis, although among patients with positive cultures who received appropriate antibiotics, there was lower mortality in patients who were given glucocorticoids (13.9%) than in the placebo group (21.7%)(P = 0.32).

**Duration of Sepsis Prior to Treatment.** When are patients most likely to benefit from a sepsis drug? How late after the onset of sepsis can a drug be expected to work?

In the study by Kreger et al. (3), in patients without "rapidly fatal" underlying disease, 40-50% of the deaths occurred in the first day after onset of bacteremia. In the HA-1A study, 13% of the placebo patients with gram-negative bacteremia died in the first 48 hours after infusion of the study material (15); this was 27% of the total deaths in this group. These data suggest that sepsis therapeutics should be most beneficial when given early in the septic response.

One would imagine that the efficacy of a drug that neutralizes endotoxin in plasma, such as an anti-endotoxin IgM antibody, would be limited to the time period during which endotoxin is circulating in the patient's blood. The duration of endotoxemia has been studied in patients with meningococcemia (26). Meningococcal LPS, measured by the Limulus lysate test, disappeared from the circulation with a half-life of 1-3 hours, then, after 4-6 hours, the half-life lengthened to 4-9 hours. The duration of endotoxemia was related to the concentration of endotoxin detected prior to therapy; although positive tests could be obtained as long as 48 hours after starting therapy, the levels of circulating LPS are much higher in patients with fulminant meningococcemia than in most other forms of gram-negative bacteremia. In Danner's study at the NIH (27), endotoxemia was detected intermittently in most patients over the 24 hours after initiating antimicrobial therapy, and some patients were positive for the first time late in this period. Endotoxin levels increased in some patients after therapy was begun, peaking at 4 hours after the administration of antimicrobials.

One might therefore expect an anti-endotoxin drug to have some benefit for at least 24 hours after the onset of sepsis. Beyond this time period, a drug intended to neutralize endotoxin might not be expected to work, since there would presumably be little circulating endotoxin to neutralize. Quite possibly, drugs that interfere with cytokine action may be beneficial when given later in the course of sepsis; this needs to be determined during the clinical evaluation of each drug.

A different problem arises when the placebo and treatment groups differ substantially in the time between the onset of sepsis and the administration of the study material. If many patients who die do so within the first 24 hours of onset of bacteremia, the longer a septic patient lives, the more likely he or she may be to survive the septic episode (or, conversely, the sicker he or she may be getting). If the study drug and

placebo are given at different times, the results may be very difficult to evaluate. It may also be difficult to determine the efficacy of the study drug according to the time interval following the onset of sepsis--a matter of great importance to the clinicians who want to use the drug most effectively.

How long should it take to administer a study drug? The mean intervals from diagnosis of sepsis to administration of study materials in the different sepsis trials are shown in the following Table:

Table 10.

| Clinical Trial | Mean interval from diagnosis of sepsis or septic shock to administration of study drug |
|---|---|
| J5 antiserum (8) | not stated |
| Corticosteroid-septic shock (9) | 12 - 27 h (group means) |
| VA Coop Glucocorticoid (11) | 2.8 h |
| Bone - Methylprednisolone (10) | 2 h |
| J5-IVIG (septic shock)(12) | 12 h (median) |
| E5(4) | 12 h or less |
| HA-1A (5) | 9.3 h (placebo), 14.3 (HA-1A) (medians) |

It is not obvious why the administration of antibody preparations should take so much longer, since both glucocorticoids and the antibody preparations were available in the pharmacies at the study hospitals.

**Laboratory parameters.** In the placebo group of the E5 study, DIC was the covariate that most strongly predicted a fatal outcome (6). Other studies have reported that ARDS worsens the prognosis in septic patients (28); hypotension and thrombocytopenia were the best predictors of impending ARDS. Interestingly, in one careful study of patients with septic shock, blood cytokine levels (TNF$\alpha$, IL-1$\beta$) were not as prognostically valuable as simple clinical parameters (underlying disease, age, arterial pH , urine output) (29). Blood lactate has some predictive value for death in patients in septic shock (30). Measurable endotoxemia, even when first detected as long as 20 hours after the initiation of therapy, is associated with severe manifestations of sepsis but not necessarily with increased mortality (27).

None of these laboratory parameters has been incorporated formally into an study protocol, except as covariates to be included in a multivariate analysis. It would be very helpful to clinicians to know if certain lab values--such as the platelet count or lactate

level--identify patients who are likely or unlikely to benefit from a given drug.

**Secondary (confirmatory) outcome measures**

It is possible to find supportive evidence for a beneficial effect in a subgroup by looking at the data for similar trends in other subgroups, for consistency in the result observed at different study sites, and for consistency over time. Non-mortality outcome measures can include discharge from the ICU, requirement for vasopressors, resolution of organ failure, discharge from the hospital, etc. If performed on the same subgroup that showed an effect on survival, the secondary analyses may be a rehash of the obvious: it is unlikely that an effect on mortality would not also be observed in, say, rates of discharge from the ICU or hospital. Secondary analyses may also represent an attempt to find positive results--such as a reduction in some morbidity in the absence of a survival benefit--but these analyses are clearly exploratory and require confirmation in another trial.

## The Choice of Placebo

The monoclonal antibody trials have been criticized for not using an indifferent monoclonal antibody as the placebo. It could be argued that the observed behavior of the monoclonal antibodies may in fact be due to some minor contaminant introduced during the production process. Indeed, prior to the HA-1A trial Dr. Ziegler proposed (31) that "the best control is an unrelated monoclonal antibody processed by the same method as the test antibody." Unfortunately, two practical considerations strongly favor using human albumin or another readily available control: (1) using another monoclonal antibody creates problems related to its specificity, the possibility of unpredictable cross-reactions, and other possible side effects; and (2) the antibody production process is very expensive--it may be unreasonable to require a company to prepare <u>two</u> monoclonal antibodies.

It has also been suggested (32) that the anti-endotoxin monoclonal antibodies should be compared to human antiserum to J5 <u>E. coli</u>, the preparation used in the successful clinical trial (8) that provided the impetus for developing the monoclonal antibodies. In effect, this would be a comparison of two unproven agents and is not recommended.

## Subgroup analysis: the HA-1A, E5, and VA glucocorticoid trials

A central problem in the evaluation of a clinical trial occurs when a drug has an apparently beneficial effect in a particular subgroup of patients, but not in the overall population studied. This situation arose in the E5, HA-1A, and VA Cooperative studies.

Oxman and Guyatt (33) have recently provided "A Consumer's Guide to Subgroup Analyses" to help clinicians evaluate trials in which an effect was seen only in one or a few subgroups of patients. Their approach asks several questions: answers are given for the HA-1A and E5 trials.

1. Was the magnitude of the difference clinically important?

   Yes, if one considers only patients in specific subgroups . On an "intention to treat" basis, however, there was no difference between the placebo and mAb groups in either trial. In the GNB subgroup of the HA-1A trial, at 28 days there were 45 deaths among the 95 patients in the placebo group (47%) and 32 deaths among the 105 HA-1A recipients (30%)(5). In the 201 patients with non-bacteremic gram-negative infections, however, the 28 day septic mortality was 38/112 (35%) in the placebo group and 37/89 (42%) in the HA-1A group ($p = 0.27$) (14). In the E5 study, there was a significant reduction in mortality in patients with gram-negative sepsis who were not in shock at the time of entry, whether or not they had gram-negative bacteremia (43% mortality in placebo group, 23% mortality in E5 group). In the total population treated, again the difference between the groups was not significant (4).

2. Was the difference statistically significant?

   In these subgroups, yes. In the HA-1A study, statistical significance was said to require a p value below 0.03 (15). Presumably a similar criterion would apply to the E5 study.

3. Did the hypothesis precede rather than follow the analysis?

   The HA-1A treatment hypothesis, as stated in the Analytical Plan, was that HA-1A would prevent **septic** deaths in patients with gram-negative **sepsis** (15). It was not anticipated that the drug would work only for preventing **all-cause** mortality in patients with gram-negative **bacteremia**. Nevertheless, the Efficacy Subgroups were specified in advance, including the all-cause, gram-negative bacteremia group. Although less information is available about the design of the E5 study, one would presume that the shock/no shock categories were specified in advance and that the antibody would be expected to have its impact in patients with early sepsis, before the onset of shock.

4. Was the subgroup analysis one of a small number of hypotheses tested?

   No. Many subgroups were analyzed in the HA-1A study. The advertised result was found in only one of these subgroups. The number of

subgroups analyzed in the E5 study is not certain.

5. Was the difference suggested by comparisons within rather than between studies?

   Yes, in both cases.

6. Was the difference consistent across studies?

   E5 was subjected to a second large, multicenter, placebo-controlled trial to test the hypothesis that it prevented death in patients with gram-negative sepsis who were not in refractory shock at the time the drug was given. In the second study, there was no difference in mortality in the placebo and E5 groups (34).

   Only one placebo-controlled HA-1A trial has been performed. The HA-1A investigators point to similarities between the results of the HA-1A trial and the previous trial of J5 antiserum (35), but the materials infused in these trials were clearly different, the entry criteria for the two trials were also different, and the basis for the reported efficacy of J5 antiserum remains unproven and highly controversial (36).

   Although E5 and HA-1A would presumably have similar mechanisms of action (both were said to bind lipid A and to protect animals from endotoxic death), they had strikingly different clinical results. E5 was said to protect patients who were not in refractory shock, whether or not they were bacteremic, whereas HA-1A protected patients with only gram-negative bacteremia who were in shock. It has been very difficult to explain this discrepancy.

7. Is there indirect evidence that supports the hypothesized difference?

   Here Oxman and Guyatt refer to studies on the biologic basis for a drug's clinical effect, or to evidence from intermediary outcomes. HA-1A was said to be associated with more rapid resolution of complicating conditions (such as ARDS) and more rapid exit from the ICU, for example, in patients with gram-negative bacteremia. It should be noted that this would be intermediary evidence supporting the efficacy of HA-1A, whereas for E5, which does not reproducibly reduce mortality, data regarding "reversal of major morbidities," reducing ICU stays, etc, are the primary basis for judging the drug and present an even more complicated challenge for subgroup analysis.

   Regarding the biological basis for the antibodies, there is considerable

controversy over the binding and protective properties of HA-1A (36,37)(see below). Although less attention has been paid to this feature of E5, it has been tested in only two or three published studies with generally unexciting results (37).

The VA Systemic Sepsis Cooperative Study Group trial of glucocorticoid therapy in septic patients with a normal sensorium provides another example of subgroup analysis (11). Here the subgroups were clearly described in the published paper: (1) patients with evidence of sepsis, (2) those with gram-negative bacteremia, and (3) those with gram-negative infections at base line. All-cause mortality at 14 days was the end point for each subgroup. In addition, the statistical importance of multiple subgroups was appreciated: "to control for multiplicity, P values (two-sided) less than 0.01 were considered to be significant..." The results are shown in Table 6 on page 11.

This trial was stopped when the sequential analysis of mortality in all patients with evidence of sepsis did not show a significant difference. Note the subgroup with gram-negative bacteremia, however: in this subgroup, glucocorticoid therapy was associated with a 75% reduction in mortality that, with relatively small numbers of patients, had the same degree of statistical significance (P value) at 14 days as did HA-1A, which was associated with a 40% reduction in mortality with much larger study groups. Intermediary evidence for glucocorticoid efficacy was found in two secondary analyses in the gram-negative bacteremia group: reductions in ARDS ($p = 0.003$) and coma ($p = 0.03$). It is also noteworthy that the glucocorticoid-treated patients in this study were more likely than the patients who received placebo to have hypotension (51% glucocorticoid, 39% placebo) and thrombocytopenia (16% vs 8%), two risk factors for death in sepsis.

The VA investigators concluded that glucocorticoid therapy was not beneficial for patients with sepsis and a normal sensorium. Regarding the gram-negative bacteremia subgroup, they commented that "although our data suggest that glucocorticoids may be efficacious in patients with sepsis from specific causative organisms, the sample size is too small and the power insufficient to permit any statistically significant or clinically meaningful conclusions to support this hypothesis." While applauding this careful conclusion and the authors' forthright presentation of the study design and analysis, one wonders what a larger study using the same clinical definitions and glucocorticoid dose would show. Perhaps the VA entry criteria, which selected patients with less severe sepsis, identified a group of patients who could benefit from this relatively inexpensive and safe intervention. (If the 75% reduction in mortality in the gram-negative bacteremia group were reproducible, one life would be saved for every 20 or so patients treated using the VA study entry criteria.)

In fact, the results of the VA glucocorticoid trial are also entirely consistent with the current understanding of the glucocorticoid effect. The drug seemed to be effective in patients with **gram-negative** bacteremia who were **treated early** with a **moderate dose** of glucocorticoid. Glucocorticoids have only been shown to protect animals from gram-

negative bacterial sepsis (or endotoxin); the requirement for early administration (to block cytokine production) was established by Beutler (38) and others; and the possibly detrimental impact of very high doses of glucocorticoid on the outcome of sepsis in animals was reported by Greisman (39).

In Oxman and Guyatt's format, questions 1, 3, 4, 5, and 7 could therefore be answered affirmatively for glucocorticoids in the VA trial. The observed positive effect was not statistically significant (question 2), and the same effect was not shown in other studies (question 6), although none of the other controlled, randomized glucocorticoid trials studied patients with early sepsis syndrome (9,10).

Table 11.

**Subgroup analysis: a comparison of three major studies**

| | HA-1A | E5 | VA steroid |
|---|---|---|---|
| 1. Magnitude of the difference | 40% decrease in mortality | 46% decrease in mortality | 75% decrease in mortality |
| 2. Statistical significance | 0.12 (14 day)<br>0.014 (28 day) | 0.01 (28 day) | 0.11 (14 day) |
| 3. A priori hypothesis | yes (2nd) | yes, probably | yes |
| 4. Small number of hypotheses | no | no | yes |
| 5. Within-study comparisons | yes | yes | yes |
| 6. Consistency across studies | not done | no | no comparable study |
| 7. Indirect evidence<br>a) Secondary outcomes: | yes | yes | yes |
| b) Basic science: | controversial | weak | supportive |

This analysis suggests that there are strong and weak features of each of the studies. None of the drugs clearly meets or fails all of the criteria, and one's conclusion is ultimately determined by the attitude one takes toward the question: should the criteria used to determine the efficacy of a drug be strict or lenient? More on this below.

## Other Issues

### Quality Control: It Should Be Concurrent

The F.D.A. may require that each lot of drug must be tested to be sure that it is the same as other lots. A strong case can be made that the criteria for quality control must be established before the clinical trial and, in most instances, the same criteria should be continued afterwards.

There are two general kinds of quality control. First, a chemical assay, such as peptide mapping or protein sequencing (for an antibody or other protein), may be used to show that the drug is consistently produced. Second, a functional assay is used to demonstrate that certain presumed protection-related properties of the drug are maintained. For example, if a drug can prevent the release of TNF$\alpha$ induced by endotoxin or bacteria in whole human blood in vitro (a property of BPI (40) and some lipid A analogs (41)), this would be the basis for a simple lot-release assay. Binding an antibody to its target epitope with the requisite specificity and affinity would be another. Protection in an animal challenge model would be still another.

This is not always a simple matter. For HA-1A, the scientific criteria used to choose the antibody that eventually went to clinical trial (35), i.e., its ability to bind LPS on intact bacteria and to protect animals from various endotoxic challenges, were endorsed in the published account of the clinical trial (5) yet, as discussed at the F.D.A. open hearing, the purified monoclonal evidently binds very poorly to smooth LPS and its ability to protect in animal models can not be reproduced reliably (15,42,43). According to a company representative, no functional assay was used to evaluate the different lots of the antibody that were used during the clinical trial (15), and although there was evidently some surveillance of these lots using chemical assays, the time at which the antibody's functional behavior changed so dramatically is not known. Moreover, there is now no functional assay that has a track record of success in a clinical trial--i.e., brand new assays would be used to evaluate the lots of the antibody to be marketed, and their relationship to the putative protective properties of the antibody may never be known.

Lesson learned: the lot-release assay for a given product should be developed and approved <u>prior</u> to the trial and used throughout it.

**The binding and protective properties of HA-1A: a comparison of different accounts**

| | **Teng et al., 1985** (35) | **Ziegler et al., 1991** (5) | **F.D.A. hearing, 1991** (15) |
|---|---|---|---|
| **Binding** | the monoclonal antibody binds to numerous smooth LPSs in ELISA assays, to smooth bacteria, "strongly" to chitin | "HA-1A has been shown to bind specifically to many endotoxins and to a broad range of clinical isolates of gram-negative bacteria."<br>"The human monoclonal IgM of Teng et al. exhibits similar cross-reactivity with a wide variety of heterologous gram-negative bacteria and their endotoxins...HA-1A, the antibody purified by Centocor, behaves identically in vitro." | Poor binding to smooth LPSs in ELISA assays, or to smooth bacteria. Binding to smooth bacteria may be increased by prior treatment with antimicrobials.<br><br>Binding to lipid A, to rough LPS, and to some smooth LPSs in a fluid-phase rate nephelometry assay.<br><br>Weak binding to chitin. |
| **Protection in animal models** | significant protection of mice from lethal bacteremia, rabbits from dermal Shwartzman reaction.<br><br>(In later studies, HA-1A protected neutropenic rabbits from pseudomonas bacteremia (44).) | "In various animal models of gram-negative bacteremia and endotoxemia, the administration of HA-1A after challenge prevents the development of the dermal Shwartzman reaction and death." | No reproducible protection in various animal models, including the dermal Shwartzman reaction. "The results are not consistently reproducible over time and from laboratory to laboratory. This lack of reproducibility has troubled workers in the field of anti-endotoxin antibodies for many years and it leads us to the conclusion that these models would not be reliable as routine potency and release assays." |

**Supervising the trial: the independent monitoring agency**

The HA-1A trial enlisted an independent coordinating center (Maryland Medical Research Institute, Baltimore) to create the randomization code, label vials, monitor compliance with blinding, audit the data, conduct interim analyses, and the like (5). This strategy may reduce the introduction of bias at various stages of the trial.

**Interim analysis.** During the performance of the study, the data are usually analyzed periodically to determine whether the study should be continued. In particular, it is important to monitor adverse events (toxicity) and, if the drug is working dramatically well, to consider stopping the trial before the planned stopping rule takes effect. These analyses should be performed by an independent group using predetermined criteria, and the results of the interim analyses should not be used to modify the study design or analytical plan. Further, if it is decided to continue the study beyond the original stopping rule in order to collect more patients in a particular group, a penalty should be paid in the analysis of statistical significance.

Problems with the interim analysis of the HA-1A trial evidently figured prominently in the recent F.D.A. decision not to license HA-1A:

"According to the FDA, the company submitted a plan to analyze the effectiveness of Centoxin over a 14 day period in specific patient groups. After gathering preliminary data, Centocor extended the review to a 28-day period and changed the patient groups that would be examined."...."An FDA source said the concern is that a firm could revise its analysis to fit the data if that procedure is allowed." (Sandra Sugawara, The Washington Post, April 16, 1992)

**Publishing the Results: What Should Journals Require?**

Oxman and Guyatt (33): "When they report the results of subgroup analyses, authors should make clear to readers how many comparisons were made and how it was decided which ones to report. Given current publication practices, however, were the reader simply to conclude that a reported interaction is real just because it is large, he or she would be wrong more often than right."

Others (21) have pointed out the common problems in published studies of clinical trials: too many end points, overuse of statistical significance testing, not enough information about study design (e.g., whether primary comparison groups and covariates were specified in advance), skimpy description of interim analyses, failure to give confidence intervals, failing to mention the determinants of trial size and the stopping rules, selecting only positive results for the summary.

Dr. Jay P. Siegel, the F.D.A. official who analyzed the HA-1A study, in a letter to the editor of the New England Journal of Medicine that was published on November 8, 1990 (45):

> "As a research scientist and clinician employed by the Food and Drug Administration, I review many clinical-trial protocols. The results of some of these trials subsequently appear in the medical literature. On occasion, the published description of the study may differ from the prospective protocol in important aspects of study design of statistical analysis--e.g., study size, clinical end-points, and statistical tests used. The potential for misuse of statistical analysis and misrepresentation of data when key study or analysis parameters are selected or modified retrospectively is tremendous, and generally such practices cannot be detected by reviewers or readers of the study report. Thus, data that are not convincing when the prospective analytical plan is applied may be "improved" by a decision to study a few more patients, report on only a definable subgroup of subjects for which the data are more convincing, use a different index of organ function or quality of life, apply a different statistical test, and so on. In other cases, such elements of the protocol were not specified in advance, and their retrospective selection is biased toward those that present the data in the "best" light. Although many retrospective decisions are not inherently improper, they must be made known to reviewers of the data to allow appropriate evaluation of the statistical inferences.
>
> To address this problem, I would suggest that medical journals recommend or require that prospective clinical-trial protocols (the whole, or key parts) be submitted along with manuscripts describing results. Since most clinical trials require approval by an institutional review boards, prospective protocols should be available. This policy would help ensure appropriate presentation and analysis of data by investigators and improved evaluation of data by reviewers. In addition, it might be expected to promote more careful consideration of study design and statistical analysis before the initiation of experimentation with human subjects."

A response to this letter was written by Dr. Arnold Relman, then the Editor of the New England Journal of Medicine:

> "In my opinion, there is no need for authors to submit their protocols as a separate supporting document. Reports of clinical trials are supposed to describe the initial protocol adequately and mention any subsequent modifications. Failure to do so constitutes a breach of scientific conduct. The rare scientists who might wish to deceive editors and readers would do so, whether or not they were asked to submit their original protocols. Given the full facts, however, editors and reviewers should be able to decide for themselves whether the quality of a study has been compromised by retrospective changes in design or analysis."

Of the various clinical trials published in the New England Journal of Medicine, only

the study by the V.A. Systemic Sepsis Cooperative Study Group (11) and the Bone methylprednisolone study (10) gave a careful description of the patient subgroups analyzed; the VA and HA-1A studies described the determinants of trial size; the VA and Bone studies described their stopping rules.

Only when readers, reviewers and editors insist on higher standards will this situation improve. Journals could lead by requiring authors of papers describing clinical trials to submit the study protocol (analytical plan), as suggested by Dr. Siegel. Readers and reviewers might insist on knowing the answers to questions such as the following:

Are the clinical definitions clear and precise?
Are the comparison groups and end points described?
Were these groups and the covariates specified in advance?
How was the size of the study population determined? Why did they stop the trial?
Was interim analysis performed? How many times? By whom?
Are time trends shown to support the important conclusions?
Are the criteria for statistical significance adjusted appropriately for the number of subgroups analyzed?
Was there an independent monitoring agency? Is it reputable?
Are confidence intervals given?
Are the results given for all the subgroups analyzed, or only for the group with the positive finding?

**Is a single placebo-controlled trial ever sufficient? The confirmatory study: who pays?**

It is very difficult to conduct clinical trials in septic patients. In particular, the striking base-line disease heterogeneity within the study populations and the variability in supportive care and antimicrobial therapy present challenges to any effort to determine the efficacy of a sepsis therapeutic. When a drug targets only a subset of the total population treated, the difficulties seem to be amplified: the smaller the subset in which efficacy is observed, the more likely one is to encounter base-line or treatment imbalances between the comparison groups. The results of the two placebo-controlled E5 studies convincingly make the point that an observation that applies to a particular subgroup should be tested in a confirmatory trial: the first study found that E5 prevented death in patients with gram-negative sepsis who were not in refractory shock ($P = 0.01$), yet the second study showed that E5 did not prevent death in this subgroup.

A drug that produces a statistically significant ($P < 0.01$) reduction in mortality in the entire population of patients might not require a confirmatory placebo-controlled trial, provided that efficacy in the total population was anticipated in the advance plan and there were no important imbalances in the distribution of covariates between the placebo and treatment groups. On the other hand, the therapeutic effect of such a drug would

be so obvious that a confirmatory trial could probably be performed using a small number of patients. The "gold standard" should be two tests that give the same result.

Who should conduct the confirmatory trial? It has been argued that a company should not be expected to do this if their product is very expensive to produce; large clinical trials can cost millions of dollars and may take two or three years to complete. Despite these concerns, the XOMA corporation conducted two large, placebo-controlled trials of E5, establishing a valuable precedent for other drugs in this field. Another mechanism for carrying out such trials may exist in the Veterans Administration hospital system; the VA Cooperative Sepsis Study Group performed one of the best clinical trials ever done in septic patients.

What about using post-marketing surveillance of these drugs to confirm their efficacy--or using the results of "compassionate use" administration of the products? These methods should not substitute for randomized, placebo-controlled trials, since there would be no valid comparison group. Using the placebo group from another trial is flawed--the patient populations are too heterogeneous. Comparing treated patients' outcome with the risk of death predicted from the patients' own APACHE scores does not take into account the quality of supportive care and antimicrobial therapy in different institutions. It is not possible simply to determine whether the product works in an individual patient, since inexplicably rapid recovery from sepsis is not that uncommon (46). In addition, HA-1A and E5 would only be expected to save one patient for every 15 or so treated; identifying that survivor would be impossible.

**Should efficacy criteria be tight or loose?**

Reflection on the VA Cooperative trial of glucocorticoid therapy and the HA-1A and E5 monoclonal antibody trials forces a comment on the remarkably different attitudes of the investigators. The VA investigators decided not to continue their trial despite a promising early result in the group with gram-negative bacteremia (75% reduction in mortality, $p = 0.11$), statistically significant results favoring glucocorticoids in two intermediary categories (coma, ARDS), an imbalance in baseline patient risk factors that favored the placebo group, and a tendency toward improved survival in all patients who received appropriate antimicrobial therapy if they also received glucocorticoids. Their analytical plan indicated that a statistically significant difference would have $P = 0.01$ or lower and that the duration of the trial would be determined by the results in the total population studied. To the VA investigators, it was extremely important to be sure that glucocorticoids really worked on an intention to treat (all-comers) basis--very stringent criteria were applied and glucocorticoids failed.

The HA-1A investigators, in contrast, stressed the importance of a 40% reduction in mortality in one of twelve subgroups of patients with gram-negative infection despite a lack of efficacy in the total population treated (or in the total population with gram-negative infection) and imbalances in baseline risk factors and antimicrobial usage that

favored HA-1A. The HA-1A (and E5) authors' attitude seems to have been much more optimistic than that of the VA investigators--they found a silver lining in a thundercloud and energetically promoted it.

Which of these attitudes best serves septic patients? In the VA case, it seems quite possible that a useful therapy was rejected--that a more optimistic outlook might have led the investigators to extend the trial to find out whether or not patients with early sepsis and gram-negative bacteremia actually benefit from glucocorticoid therapy. The HA-1A (and E5) approach maximizes the opportunity to discover that a drug is beneficial, yet it also increases the likelihood that an ineffective drug will be widely used. We probably want to be somewhere between these extremes: to minimize the risk of both Type I and Type II errors.

## Suggestions for Future Studies

Much of the following could be worked out by experienced investigators in academia, industry, and the F.D.A. A consensus approach to future studies in septic patients is badly needed.

**Standardized clinical definitions.** If all future studies used the same clinical definitions, cross-study comparisons would be much easier. The entry criteria could be tailored to the individual drug depending upon its likely effect--in patients with early vs. late sepsis, for example. It would be particularly helpful if different studies used the same definitions of shock, refractory shock, and organ failure, and if studies would record the same basic set of physiologic and laboratory data on each patient.

**Standardized outcome end points, subgroups.** It is too much to expect different investigators/companies to choose precisely the same end points. But certain features seem to work well: 14 and 28 day mortality end points; septic and all-cause mortality; carefully defined disease categories. Mortality should be analyzed according to the presence or absence of clinically useful parameters such as altered mental status, septic shock, refractory septic shock, and thrombocytopenia. A small number of primary and secondary treatment comparisons should be specified in advance.

**Concurrent quality control.** The lot-release assay should be established prior to the trial and used on the lots of drug that are given during the trial.

**Independent trial supervision.** The independent monitoring agency would ideally carry out not only the randomization, drug delivery, interim analysis, etc., but also (in collaboration with the academic investigators) analyze and submit the efficacy data to the company and the F.D.A.

**Larger studies.** If a drug targets only a subpopulation of patients with sepsis, it

may be best to test the drug in a study population that is sufficiently large to minimize the risk of imbalances in critical covariates within the targeted subgroup. Unfortunately, the results of the trials discussed today suggest that the necessary study size may be prohibitively large. Perhaps the dice are loaded against the evaluation of drugs that only benefit a subset of this extremely diverse and complex patient population.

---

ACKNOWLEDGMENTS. I thank Drs. Joan Reisch, William Baine, and Robert Haley for trying to teach me clinical epidemiology and statistics. I am also grateful to Drs. Shaw Warren (Mass. General Hospital) and Robert Danner (NIH) for helpful discussions.

Because this discussion deals in part with products that are currently undergoing evaluation by the F.D.A. and are being considered for use in hospitals, its seems appropriate to mention a possible conflict of interest. I share patents with Catherine L. Hall for the purification of acyloxyacyl hydrolase, an enzyme that may detoxify endotoxin, and for several possible uses of partially deacylated lipopolysaccharides. The patents do not cover the use of the enzyme as a therapeutic, and it seems unlikely now that either the enzyme or the deacylated LPS would be useful for treating septic patients. ZymoGenetics, Inc. (Seattle, WA) collaborated with us to clone the cDNA for the enzyme and provided funding for my laboratory prior to August, 1991. I have had no professional or financial relationship with Centocor, XOMA, or any other biotechnology company.

APPENDIX

**Can one diagnose gram-negative sepsis before culture results are available?**

Several of the new sepsis therapies aim to neutralize gram-negative bacterial endotoxin, and therefore should be useful only in patients with gram-negative bacterial sepsis. A successful method for diagnosing gram-negative sepsis might signficantly increase the cost-effectiveness of these drugs (47). The size and cost of clinical efficacy trials for these products might also be reduced, since only patients with gram-negative sepsis would be entered into the trials. Attempts to make a specific diagnosis of gram-negative bacteremia or gram-negative sepsis have focused almost entirely on the detection of endotoxin in plasma. The most widely used and successful assay is based on the ability of endotoxin to initiate a clotting cascade in Limulus polyphemus (horseshoe crab) amoebocyte lysate. A chromogenic clotting substrate may be used to obtain a quantitative read-out. Standardization of the assay is poor, and there are questions about its specificity. Nevertheless, many experts feel that it is a reasonable method for detecting, and even quantitating, endotoxin in plasma.

Published results suggest that the Limulus assay may be useful for identifying many, but not all, patients with gram-negative sepsis, and that it may sometimes be positive in patients with gram-positive bacterial or fungal infections.

TABLE 12.

**Endotoxin detection in septic patients**

| | **Endotoxemia in septic patients (No. positive/total tested)(%)** | | |
|---|---|---|---|
| **Author, year** | **Gram-negative bacteremia** | **No gram-negative bacteremia** | **Patient population** |
| Shenep (48), 1988 | 9/10 (90%) | 3/16 (19%) | children with sepsis |
| Van Deventer (49), 1988 | 16/19 (84%) | 4/10 (40%) | adults with presumed sepsis |
| Brantzaeg (26), 1989 | 24/35 (69%) | 0/7 (0%) | meningococcal disease |
| Danner (27), 1991 | 11/19 (58%) | 32/81 (39%) | adults with sepsis; many were neutropenic |

In these studies endotoxin determinations were usually done on blood specimens obtained prior to the administration of antimicrobials. The detection threshold for the assays were very low (5 - 25 pg/ml). Serial endotoxin measurements were made over the first 24 hours of ICU admission by Danner et al. They found that endotoxin was detected intermittently in many patients and that the cumulative percentage of septic patients with endotoxemia increased over time (from 20% at admission to 40% at 20 hours). In some patients, endotoxin levels rose above entry values despite the administration of appropriate antibiotics, peaking at 4 hours after the start of sample collection (27). A positive test for endotoxin identified a subset of patients with severe sepsis, yet endotoxemia occurred frequently in the absence of positive blood cultures for gram-negative bacteria.

Summary: the existing endotoxin detection methods cannot be counted upon to identify all patients with gram-negative bacterial sepsis (sensitivity), or to distinguish gram-negative sepsis from sepsis of other etiologies (specificity). A positive LAL test may identify patients with more severe septic illness (27). Although the test may be performed within a few hours, it is unlikely that this methodology would be very useful for clinical decision-making (treat or not treat). It is possible that the specificity and sensitivity of endotoxin detection may be improved by recent developments in assay design. For example, using solid-phase anti-endotoxin antibodies (or bactericidal permeability-increasing protein) to "capture" plasma endotoxin and Limulus lysate to detect the captured endotoxin greatly improves the specificity of the assay without sacrificing sensitivity (50).

1. Johnston, J. 1992. Molecular science sets it sights on septic shock. *J. NIH Res.* 3:61-65.

2. Bone, R.C. 1991. Sepsis, the sepsis syndrome, multi-organ failure: A plea for comparable definitions. *Ann. Intern. Med.* 114:332-333.

3. Kreger, B.E., D.E. Craven, and W.R. McCabe. 1980. Gram-negative bacteremia. IV. Re-evaluation of clinical features and treatment in 612 patients. *Am. J. Med.* 68:344-355.

4. Greenman, R.L., R.M.H. Schein, M.A. Martin, R.P. Wenzel, N.R. MacIntyre, G. Emmanuel, H. Chmel, R.B. Kohler, M. McCarthy, J. Plouffe, J.A. Russell, and XOMA Sepsis Study Group. 1991. A controlled clinical trial of E5 murine monoclonal IgM antibody to endotoxin in the treatment of gram-negative sepsis. *JAMA* 266:1097-1102.

5. Ziegler, E.J., C.J. Fisher,Jr., C.L. Sprung, R.C. Straube, J.C. Sadoff, G.E. Foulke, C.H. Wortel, M.P. Fink, R.P. Dellinger, N.N.H. Teng, I.E. Allen, H.J. Berger, G.L. Knatterud, A.F. LoBuglio, C.R. Smith, and HA-1A Sepsis Study Group. 1991. Treatment of gram-negative bacteremia and septic shock with HA-1A human monoclonal antibody against endotoxin--A randomized, double-blind, placebo-controlled trial. *N. Engl. J. Med.* 324:429-436.

6. Martin, M.A., R.P. Wenzel, and K.J. Gorelick. 1991. Gram-negative bacterial sepsis in hospitals in the United States--Natural history in the 1980s. *Prog. Clin. Biol. Res.* 367:111-119.

7. Sprung, C.L. 1991. Definitions of sepsis--have we reached a consensus?. *Crit. Care. Med.* 19:849-851.

8. Ziegler, E.J., J.A. McCutchan, J. Fierer, M.P. Glauser, J.C. Sadoff, H. Douglas, and A.I. Braude. 1982. Treatment of gram-negative bacteremia and shock with human antiserum to a mutant Escherichia coli. *New Eng. J. Med.* 307:1225-1230.

9. Sprung, C.L., P.V. Caralis, E.H. Marcial, M. Pierce, M.A. Gelbard, W.M. Long, R.C. Duncan, M.D. Tendler, and M. Karpf. 1984. The effects of high-dose corticosteroids in patients with septic shock. A prospective, controlled study. *New Eng. J. Med.* 311:1137-1143.

10. Bone, R.C., C.J. Fisher,Jr., T.P. Clemmer, G.J. Slotman, C.A. Metz, R.A. Balk, and and the Methylprednisolone Severe Sepsis Study Group. 1987. A controlled clinical trial of high-dose methylprednisolone in the treatment of severe sepsis and septic shock. *New Eng. J. Med.* 317:653-658.

11. The Veterans Administration Systemic Sepsis Cooperative Study Group 1987. Effect of high-dose glucocorticoid therapy on mortality in patients with clinical signs of systemic

sepsis. *New Eng. J. Med.* 317:659-665.

12. Calandra, T., M.P. Glauser, J. Schellekens, and J. Verhoef. 1988. Treatment of gram-negative septic shock with human IgG antibody to *Escherichia coli* J5: A prospective, double-blind, randomized trial. *J. Infect. Dis.* 158:312-319.

13. Sprung, C.L., P.N. Peduzzi, C.H. Shatney, R.M. Schein, M.F. Wilson, J.N. Sheagren, L.B. Hinshaw, and The Veterans Administration Systemic Sepsis Cooperative Study Group. 1990. . *Crit. Care. Med.* 18:801-805.

14. Siegel, J.P. 1991. Statistical presentation at the open meeting of the Vaccines and Related Biological Products Advisory Committee, Department of Health and Human Services, Public Health Service, Food and Drug Administration, held on Sept. 4, 1991, in Washington,D.C. (UnPub)

15. 1991. Transcript of the open meeting of the Vaccines and Related Biological Products Advisory Committee, Department of Health and Human Services, Public Health Service, Food and Drug Administration, held on Sept. 4, 1991 in Washington, D.C. *none*

16. Tanio, C.P., and H.I. Feldman. 1991. The HA-1A monoclonal antibody for gram-negative sepsis (Letter). *New Eng. J. Med.* 325:280.

17. McCabe, W.R., and G.G. Jackson. 1962. Gram-negative bacteremia I. Etiology and ecology. *Arch. Int. Med.* 110:83-91.

18. Knaus, W.A., E.A. Draper, D.P. Wagner, and J.E. Zimmerman. 1985. APACHE II: A severity of disease classification system. *Crit. Care. Med.* 13:818-829.

19. Knaus, W.A., D.P. Wagner, E.A> Draper, J.E. Zimmerman, M. Bergner, P.G. Bastos, C.A. Sirio, D.J. Murphy, T. Lotring, A. damiano, and F.E.Jr. Harrell. 1991. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 100:1619-1636.

20. Knaus, W.A., D.P. Wagner, and J. Lynn. 1991. Short-term mortality predictions for critically ill hospitalized adults: science and ethics. *Science* 254:389-394.

21. Pocock, S.J., M.D. Hughes, and R.J. Lee. 1987. Statistical problems in the reporting of clinical trials. *New Eng. J. Med.* 317:426-432.

22. Weil, M.H., H. Shubin, and M. Biddle. 1964. Shock caused by gram-negative microorganisms. Analysis of 169 cases. *Ann. Intern. Med.* 60:384-400.

23. McCue, J.D. 1985. Improved mortality in gram-negative bacillary bacteremia. *Arch. Int. Med.* 145:1212-1216.

24. Bryan, C.S., K.L. Reynolds, and E.R. Brenner. 1983. Analysis of 1,186 episodes of gram-negative bacteremia in non-university hospitals: the effects of antimicrobial therapy. *Rev. Infect. Dis.* 4:629-638.

25. Bogart, W.C.Jr., and S.A. Siegel. 1991. The human monoclonal antibody HA-1A: studies on the epitope location within the endotoxin molecule and epitopic exposure on the surface of viable gram-negative bacteria. *Circ. Shock* 34:119.(Abstr.)

26. Brandtzaeg, P., P. Kierulf, P. Gaustad, A. Skulberg, J.N. Bruun, S. Halvorsen, and E. Sorensen. 1989. Plasma endotoxin as a predictor of multiple organ failure and death in systemic meningococcal disease. *J. Infect. Dis.* 159:195-204.

27. Danner, R.L., R.J. Elin, J.M. Hosseini, R.A. Wesley, J.M. Reilly, and J.E. Parillo. 1991. Endotoxemia in human septic shock. *Chest* 99:169-175.

28. Fein, A.M., M. Lippmann, H. Holtzman, A. Eliraz, and S.K. Goldberg. 1983. The risk factors, incidence, and prognosis of ARDS following septicemia. *Chest* 83:40-42.

29. Calandra, T., J.-D. Baumgartner, G.E. Grau, M.-M. Wu, P.-H. Lambert, J. Schellekens, J. Verhoef, M.P. Blauser, and Swiss-Dutch J5 Immunoglob Study Grp. 1990. Prognostic values of tumor necrosis factor/cachectin, interleukin- 1, interferon-$\alpha$, and interferon-gamma in the serum of patients with septic shock. *J. Infect. Dis.* 161:982-987.

30. Bakker, J., M. Coffernils, M. Leon, P. Gris, and J.-L. Vincent. 1991. Blood lactate levels are superior to oxygen-derived variables in predicting outcome in human septic shock. *Chest* 99:956-962.

31. Ziegler, E.J. 1988. Protective antibody to endotoxin core: the emperor's new clothes?. *J. Infect. Dis.* 158:286-290.

32. Wolff, S.M. 1991. Monoclonal antibodies and the treatment of gram-negative bacteremia and shock. *New Eng. J. Med.* 324:486-487.

33. Oxman, A.D., and G.H. Guyatt. 1992. A consumer's guide to subgroup analyses. *Ann. Intern. Med.* 116:78-84.

34. Wenzel, R., R. Bone, A. Fein, R. Quenzer, J. Schentag, K.J. Gorelick, N.I. Wedel, and T. Perl. 1991. Results of a second double-blind, randomized, controlled trial of antiendotoxin antibody E5 in gram-negative sepsis. Abstracts of the Interscience Conference on Antimicrobial Agents and Chemotherapy, Chicago, Illinois, Sept. 29 - Oct. 2, 1991, extended abstract #1170, American Society for Microbiology.

35. Teng, N.N.H., H.S. Kaplan, J.M. Hebert, C. Moore, H. Douglas, A. Wunderlich, and

A.I. Braude. 1985. Protection against gram-negative bacteremia and endotoxemia with human monoclonal IgM antibodies. *Proc. Natl. Acad. Sci. USA* 82:1790-1794.

36. Zanetti, G., M.P. Glauser, and J.D. Baumgartner. 1991. Use of immunoglobulins in prevention and treatment of infection in critically ill patients: review and critique. *Rev. Infect. Dis.* 13:985-992.

37. Warren, H.S., R.L. Danner, and R.S. Munford. 1992. Anti-endotoxin monoclonal antibodies. *New Eng. J. Med.*, in press (April 23, 1992)

38. Beutler, B., N. Krochin, I.W. Milsark, C. Luedke, and A. Cerami. 1986. Control of cachectin (tumor necrosis factor) synthesis: mechanisms of endotoxin resistance. *Science* 232:977-980.

39. Greisman, S.G. 1982. Experimental gram-negative bacterial sepsis: optimal methylprednisolone requirements for prevention of mortality not preventable by antibiotics alone. *Proc. Soc. Exp. Biol. Med.* 170:436-442.

40. Eng Ooi, C., J. Weiss, M.E. Doerfler, and P. Elsbach. 1991. Endotoxin-neutralizing properties of the 25 kD N-terminal fragment and a newly isolated 30 kD C-terminal fragment of the 55-60 kD bactericidal/ permeability-increasing protein of human neutrophils. *J. Exp. Med.* 174:649-655.

41. Kovach, N.L., E. Yee, R.S. Munford, C.R.H. Raetz, and J.M. Harlan. 1990. Lipid IVA inhibits synthesis and release of tumor necrosis factor induced by LPS in human whole blood ex vivo. *J. Exp. Med.* 172:77-84.

42. Baumgartner, J.D., D. Heumann, J. Gerain, P. Weinbreck, G.E. Grau, and M.P. Glauser. 1990. Association between protective efficacy of anti-lipopolysaccharide (LPS) antibodies and suppression of LPS-induced tumor necrosis factor $\alpha$ and interleukin 6. Comparison of O side chain- specific antibodies with core LPS antibodies. *J. Exp. Med.* 171:889-896.

43. Baumgartner, J.D. 1990. Monoclonal anti-endotoxin antibodies for the treatment of gram-negative bacteremia and septic shock. *Eur. J. Clin. Microbiol. Infect. Dis.* 9:711-716.

44. Ziegler, E.J., N.N.H. Teng, H. Douglas, A. Wunderlich, H.J. Berger, and S.D. Bolmer. 1987. Treatment of Pseudomonas bacteremia in neutropenic rabbits with human monoclonal IgM antibody against E. coli lipid A. *Clin. Res.* 35:619A (Abstr.)

45. Siegel, J.P. 1990. Editorial review of protocols for clinical trials. *New Eng. J. Med.* 323:1355.

46. DeMaria, A., D.E. Craven, J.J. Heffernan, and W.R. McCabe. 1985. Naloxone versus

placebo in treatment of septic shock. *Lancet* 1:1363-1365.

47. Schulman, K.A., H.A. Glick, H. Rubin, and J.M. Eisenberg. 1991. Cost-effectiveness of HA-1A monoclonal antibody for gram- negative sepsis. Economic assessment of a new therapeutic agent. *JAMA* 266:3466-3471.

48. Shenep, J.L., P.M. Flynn, F.F. Barrett, G.L. Stidham, and D.F. Westenkirchner. 1988. Serial quantitation of endotoxemia and bacteremia during therapy for gram-negative bacterial sepsis. *J. Infect. Dis.* 157:565-568.

49. van Deventer, S.J.H., H.R. Buller, J.W. ten Cate, A. Sturk, and W. Pauw. 1988. Endotoxaemia:an early predictor of septicaemia in febrile patients. *Lancet* 1:605-609.

50. Mertsola, J., L. Cope, R.S. Munford, G.H. McCracken,Jr., and E.J. Hansen. 1991. Detection of experimental Haemophilus influenzae type b bacteremia and endotoxemia by means of an immunolimulus assay. *J. Infect. Dis.* 164:353-358.