Relationship of Systemic Inflammatory Response Syndrome to Organ Dysfunction, Length of Stay, and Mortality in Critical Surgical Illness

Effect of Intensive Care Unit Resuscitation

Mia Talmor, MD; Lynn Hydo, BS; Philip S. Barie, MD, FCCM

Background: A systemic proinflammatory response has been implicated in the pathogenesis of organ dysfunction. The effects of surgery, surgical stress, anesthesia, and subsequent intensive care unit (ICU) resuscitation may affect the components of the systemic inflammatory response syndrome (SIRS) score (temperature, heart rate, respiratory rate, and white blood cell count). Any SIRS scores calculated within 24 hours after surgery or at the onset of nonoperative resuscitation may overestimate the proinflammatory response itself, making quantification of SIRS at that time potentially too sensitive. We hypothesized that SIRS attributable to ICU resuscitation can be quantitated, and that SIRS after the first day of therapy in the ICU correlates with several outcomes.

Methods: Prospective analysis of 2300 surgical ICU admissions during a 49-month period. Acute Physiology and Chronic Health Evaluation III (APACHE III) scores were recorded after 24 hours. Daily and cumulative multiple organ dysfunction scores (0-4 points for each of 6 organs, 24 points total) and SIRS scores (1 point for each parameter, 4 points total) were recorded. Defined end points were hospital mortality, days in the ICU, and organ dysfunction.

Results: On day 1, 49.4% of patients had SIRS (score ≥2), whereas 34.5% of patients who remained in the ICU had SIRS (score ≥2) on day 2 (P<.001). The SIRS score decreased by a mean of 0.8 points from day 1 to day 2, regardless of the type of admission. A SIRS score that decreased on day 2, in comparison with the score on day 1, resulted in less mortality than a unchanged or higher score on day 2 (11% vs 18% vs 22%, P<.001). Systemic inflammatory response scores were higher for nonsurvivors than survivors on each of the first 7 days in the ICU. The day 2 SIRS score correlated well with the admission APACHE III score (P<.001) and all defined end points (all P<.001). The day 2 SIRS score also correlated with the day 2 multiple organ dysfunction score (P<.001). By multiple logistic regression, APACHE III (P<.001), day 2 SIRS score (P<.01) (but not day 1 SIRS score, P=.99), and day 2 multiple organ dysfunction score (P<.001) (but not day 1 multiple organ dysfunction score, P=.81) predicted mortality.

Conclusions: Systemic inflammatory response syndrome attributable to surgery or surgical stress can be quantitated. Twenty-four hours of ICU resuscitation results in a decline in the SIRS score. The magnitude of the proinflammatory response on the second ICU day may be a useful predictor of outcome in critical surgical illness.

Arch Surg. 1999;134:81-87
PATIENTS AND METHODS

This study was conducted during a 49-month period in the general surgical ICU of an urban university tertiary care medical center that provides surgical care to its local community and to tertiary care referrals, and functions also as a level I trauma center. Although the institution serves as a regional center for burn care and cardiothoracic surgery, those patients are cared for in separate, dedicated units and thus are excluded from this analysis.

All variables analyzed in this study were calculated prospectively, with the exception of SIRS scores, which were calculated retrospectively from raw data collected prospectively. Acute Physiology and Chronic Health Evaluation II and III (APACHE II and APACHE III) scores were calculated manually as described in the original reports and reported for our patient population previously. The worst physiologic values during the first 24 hours of critical care were used for calculation of the APACHE scores. All patients were analyzed until hospital discharge or death. In circumstances where a patient was readmitted to the surgical ICU during the same hospital stay, all admissions were included for the ICU length of stay analysis, but hospital mortality was calculated as a percentage of patients.

The MOD score of Marshall et al was calculated daily and a maximal score was also determined. The MOD score grades increasing severity of dysfunction of 6 organ systems (cardiac, central nervous system, hematologic, hepatic, pulmonary, and renal) on a 4-point scale, so that a maximum of 24 points can be accrued. The maximal MOD score was calculated by summing the individual worst values at any time during the ICU stay for each of the 6 organ systems described. Therefore, the maximal MOD score could only increase for the duration of observation, whereas the daily MOD score would fluctuate with clinical improvement or deterioration. Patients were considered to have some degree of organ dysfunction if the MOD score was 1 point or higher. This system was compared previously with that of Knaus and Wagner, which considers organ failure to be an all-or-nothing phenomenon, and was found to correlate closely.

The SIRS scores were calculated similarly. The SIRS score assigns 1 point for each parameter (temperature, white blood cell count, heart rate, and respiratory rate), so that a maximum of 4 points can be accrued. A maximal SIRS score was calculated by summing the individual worst value in each of the 4 parameters. Thus, a patient might have a maximal SIRS score of 4, even if the patient had scored 1 point in each different parameter at different points in the ICU admission. A daily SIRS score, which fluctuates on a day-to-day basis similar to the daily MOD score, was also calculated each day of the ICU stay.

The ICU length of stay, hospital length of stay, and maximal MOD score were assessed for each daily SIRS score group (eg, day 1 vs day 2, 3, and so on), and compared using a 1-way analysis of variance (ANOVA) with repeated measures. Differences between groups were identified using the Tukey post hoc test. Mortality rates and the incidences of MOD were compared using a x² contingency table analysis for multiple group comparisons. Comparison of group mean data for continuous variables between survivors and nonsurvivors was made using an unpaired nonparametric Wilcoxon analysis. Subset analyses were performed on survivors vs nonsurvivors, and of elective or emergent operative vs nonoperative admissions. Patients admitted electively were those who underwent a scheduled operation on the day of admission, according to the published operating room schedule. These patients were considered to be elective if the ICU admission followed a destabilizing event during surgery or in the postanesthesia care unit. Nonoperative patients were those admitted with illness or injury (eg, pancreatitis or multiple trauma) or those who sustained a superimposed insult (eg, myocardial infarction or pulmonary embolism) necessitating ICU care at least 24 hours after surgery. Multivariate assessments of the effect of independent variables on dependent outcome variables (magnitude of MOD, length of stay, and mortality) were made with a multivariate logistic regression analysis.

All statistical analyses were performed by microcomputer using commercial software (StatView 4.0.1 with SuperANOVA 1.1.1, Abacus Concepts Inc, Berkeley, Calif). Statistical significance was determined at an α level of .05. Group mean data are expressed as mean±SE.

RESULTS

Data from 2300 consecutive admissions for critical surgical illness, postoperative complications, or postoperative care were evaluated. The composition of the population within the ICU during the first week of care is depicted in Figure 1. The mean age for the group was 64.6±0.3 years, the mean APACHE II score was 13.9±0.2, and the mean APACHE III score was 44.2±0.6. Mean ICU length of stay was 5.5±0.2 days, and average hospital length of stay was 22.4±0.6 days. Forty-nine per-
cent of patients developed some degree of organ dysfunction, and hospital mortality for the entire group was 8.3%. The mean maximal SIRS score for the group was 1.9 ± 0.1. On day 1, 49.4% of patients had SIRS (≥2 criteria), whereas 34.5% of patients who remained in the ICU had SIRS on day 2 (P < .001). The SIRS scores on successive days, as a function of ultimate ICU length of stay, are depicted in Figure 2. The day 1 SIRS scores were successively higher (P < .01) for patients who were destined for longer ICU length of stay, but the rate of resolution of SIRS was similar. On average, the SIRS score on the day of discharge from the ICU was 0.5 to 0.95.

Mean daily SIRS scores are depicted (Figure 3) for survivors and nonsurvivors for the first 7 days of the ICU stay. This analysis is not based on the day of discharge, therefore patients with multiday stays were considered repetitively. The SIRS score was significantly higher (P < .01) for nonsurvivors on each day, as were age, APACHE II and III scores, ICU length of stay, and hospital length of stay (all P < .001) (data not shown). The APACHE III scores, ICU length of stay, MOD score, and mortality rate all increased with higher day 2 SIRS scores (P < .001) (Table 1). In actuality, a higher SIRS score related to higher mortality for each of the first 4 days in the ICU (P < .001) (Figure 4). Mortality was negligible for patients who never had any component of SIRS, but increased progressively (P < .01) for higher scores as the ICU stay became prolonged. Mortality was highest (38% and 45%, respectively) for patients with a SIRS score of 4 on day 2 or day 3.

Among patients who had an ICU length of stay of 2 days or more, a decreased day 2 SIRS score, compared with the same-patient SIRS score on day 1, resulted in less mortality than an unchanged or positive score (11% vs 18% vs 22%, P < .001). In contrast, the mean SIRS score of patients discharged in less than 24 hours was 1.0 ± 0.1 and mortality was 3.8%, due in part to 22 patients who died within the first 24 hours.

A subset analysis of elective or emergent operative vs nonoperative admissions was performed (Figure 5 and Figure 6). Whereas the mean SIRS score on day 1

<table>
<thead>
<tr>
<th>SIRS Score</th>
<th>Admission APACHE III Score†</th>
<th>ICU Length of Stay†</th>
<th>MOD Score†</th>
<th>Mortality, %†</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>47.6 ± 1.8</td>
<td>4.0 ± 0.4</td>
<td>3.6 ± 0.2</td>
<td>3.8</td>
</tr>
<tr>
<td>1</td>
<td>51.8 ± 1.7</td>
<td>7.0 ± 0.7</td>
<td>4.5 ± 1.2</td>
<td>8.2</td>
</tr>
<tr>
<td>2</td>
<td>60.3 ± 1.9</td>
<td>11.3 ± 1.0</td>
<td>7.0 ± 0.3</td>
<td>18.4</td>
</tr>
<tr>
<td>3</td>
<td>58.0 ± 2.3</td>
<td>12.6 ± 1.4</td>
<td>7.5 ± 0.5</td>
<td>24.2</td>
</tr>
<tr>
<td>4</td>
<td>78.5 ± 4.9</td>
<td>9.8 ± 2.1</td>
<td>8.5 ± 0.6</td>
<td>40.0</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SE unless otherwise indicated. SIRS indicates systemic inflammatory response syndrome; APACHE III, Acute Physiology and Chronic Health Evaluation III; ICU, intensive care unit; and MOD, multiple organ dysfunction.

*P < .001.

©1999 American Medical Association. All rights reserved.

Downloaded From: http://archsurg.jamanetwork.com/pdfaccess.ashx?url=/data/journals/surg/9414/ on 06/01/2017
in the elective group (1.5 ± 0.1) was lower than in the emergent or nonoperative groups (both 2.1 ± 0.1) (P < .001), the decrease in SIRS by 0.8 points from day 1 to day 2 was identical (Figure 5). From a nadir on day 2, SIRS scores increased in aggregate on each subsequent day, reflective of the illness severity that prolongs ICU length of stay. Similarly, MOD scores were higher in relation to prolonged ICU length of stay (Figure 6). The increase in SIRS and MOD scores was related to ICU length of stay, in each subset of patients, and in inverse relation to the number of patients remaining in the ICU (Figure 7).

The multivariate analysis of the data is summarized in Table 2. The APACHE III score (P = .01), day 2 SIRS score (P < .001) (but not day 1 SIRS score, P = .88) and day 2 MOD score (P < .001) (but not day 1 MOD score, P = .62) were independent predictors of mortality.

The data indicate that SIRS attributable to surgical stress and ICU resuscitation can be quantitated. Regardless of admission type, the mean SIRS score decreased by 0.8 points from day 1 to day 2 of the ICU stay, reflective of ICU resuscitation. Elevated day 1 SIRS scores are largely artifact in elective surgery patients because the mean score...
decreased by 50% to only 0.8 ± 0.1 points on day 2. Furthermore, the SIRS score decreased by an identical amount from day 1 to day 2 in emergent operative and nonoperative patients, making it plausible that the similarity of responses reflects either resuscitation to comparable end points or adherence to similar principles (eg, pain control, suppression of tachycardia),\(^5\) despite the ostensibly dissimilarity of the groups.

In 1995, Pittet et al\(^4\) investigated the epidemiology of SIRS in a population of 170 surgical ICU patients. The SIRS criteria (≥2) were met by 93% of the patients, whereas only 8.2% died. The authors concluded that because of very high sensitivity and poor specificity, SIRS did not identify those patients who died ultimately. They did not investigate the relationship of SIRS to the incidence or magnitude of MOD. In a critique of Pittet et al, Sibbald et al\(^3\) suggested that the high rate of SIRS reported by Pittet et al could be explained by differences in definitions for SIRS proposed by the American College of Chest Physicians/Society of Critical Care Medicine panelists and those used by the authors. Illustrating their point, Sibbald et al\(^3\) suggested that in postoperative patients, the American College of Chest Physicians/Society of Critical Care Medicine criteria would not have defined SIRS on ICU admission because the acute alteration from baseline occurred due to another known cause (namely, the surgery). Sibbald et al suggested further that by 12 hours postoperatively, ongoing abnormalities are referable to a baseline (ICU admission) and a diagnosis of SIRS may be confirmed once the effects of anesthesia several hours before had subsided. Menger and Vollmar\(^6\) concurred in principle, but argued that patients might be affected by major surgery for several days, and that major surgical tissue injury, as with entities such as trauma, might be an important factor in the development of SIRS and severe sepsis.

Indeed, the quantification of SIRS is potentially overly sensitive in surgical populations. Major surgical stress, anesthesia, and postoperative pain can result in a systemic response (temperature >38°C, leukocytosis, tachypnea, and tachycardia) that can mimic acute inflammation, but the response should be short-lived if resuscitation is adequate. Therefore, a SIRS score obtained within 24 hours of ICU admission is a poor predictor of outcome, as Sibbald et al\(^3\) suggested. However, contrary to the contention of Menger and Vollmar,\(^6\) this proinflammatory response seems to be pathological if it persists beyond 24 hours. In this study, an elevated SIRS score obtained between 24 and 48 hours after admission, despite aggressive resuscitation, predicted an increased mortality. Moreover, it is not only the day 2 score in isolation, but failure of the score to decrease or indeed to increase during the second 24 hours (which indicates an ongoing or superimposed proinflammatory response) that is important.

Haga et al\(^10\) retrospectively studied 292 ICU patients after gastrointestinal tract surgery. They found that not only the number of positive SIRS criteria but also the duration of SIRS correlated with an adverse outcome. These authors examined a surgical stress parameter (blood loss indexed to body mass), operative time, and serum C-reactive protein concentrations and documented that surgery itself can lead to a proinflammatory state, as had been hypothesized.\(^5,6\) However, patients who recovered from early postoperative SIRS had a lower incidence of MOD than did those in whom SIRS persisted. Specifically, SIRS (≥2 criteria) continuing consecutively for 2 days after postoperative day 3 had a higher incidence of both infectious and noninfectious postoperative complications, as well as a higher incidence of MOD. We agree that persistent SIRS is deleterious, but our data indicate that the magnitude of SIRS on the second postoperative day (3 days before the study of Haga et al) correlates strongly with adverse outcomes.

Our findings support the current theories regarding the pathogenesis of SIRS and organ dysfunction syndrome popularized by Bone.\(^27,28\) An elevated day 1 SIRS score may be secondary to a local proinflammatory coping reaction. However, by day 2, if a putative antiinflammatory response has not balanced the proinflammatory response, SIRS persists, resulting in an increased incidence of MOD and death. Our findings also support those who point to the narrow time window for effective resuscitation after surgical or traumatic stress. Abramson et al\(^19\) showed that trauma patients who were resuscitated adequately within 24 hours (measured by normalization of serum lactate levels) all survived. In contrast, only 75% of patients survived if their lactate levels normalized within 24 to 48 hours, and 14% of patients survived if their lactate levels did not normalize within 48 hours. Similarly, a decreased day 2 SIRS score (presumably reflective of adequate resuscitation and postoperative care) correlates with a significantly lower mortality than an increased or unchanged score.

The effect of 24 hours of ICU resuscitation for both operative as well as nonoperative patients is clear from our data. Whereas patients admitted emergently had higher SIRS scores than those admitted electively on day 1, the decrease in SIRS from day 1 to day 2 by 0.8 points was identical regardless of the cause for admission, reflecting the effect of consistent ICU resuscitation. For this reason, SIRS scores calculated prior to resuscitation may result in artificially elevated scores that correlate poorly with outcome.

Current hypotheses of the pathogenesis of MOD may need to be reexamined. In the classic paradigm that affects most patients, organ failure develops catastrophically as a late consequence of critical illness, as induced by a “2-hit” phenomenon.\(^22\) The 2-hit phenomenon represents priming of the inflammatory response by a primary insult (eg multiple trauma), followed by activation of the response by a secondary insult (eg, pneumonia). This study and our previous data indicate that clinically meaningful derangements are present and detectable very early in the course of critical illness.\(^8\) Sequential elements necessary for the development of SIRS or MOD may occur within hours rather than days.\(^21\) The importance of early, aggressive resuscitation is underscored, but the time window for successful intervention may be narrow.

The early identification of seriously ill patients at risk for an adverse outcome would yield several benefits. Patient triage and resource utilization will become more efficient, and accurate prognostic information would enhance the dialogue with patient or family regarding the plan of care and the expected result. A simple-to-calculate indicator that correlates closely with outcomes in groups of patients has the advantage of ease of use, but is not yet significantly robust for outcome assessments in individuals. Of equal importance is the recognition that physiologic responses so early (day 2) in the ICU stay reflect outcomes. In contrast to the long-standing observation that organ dysfunction occurs as a late, catastrophic event, meaningful derangements are detectable early, and the window for successful resuscitation appears to be narrow.

Reprints: Philip S. Barie, MD, FCCM, Department of Surgery, The New York Presbyterian Hospital–Cornell Medical Center, 525 East 68th St, New York, NY 10021.

REFERENCES


a meaningful prediction, and it might, but having not performed receiver operating characteristic curves or goodness-of-fit testing for SIRS versus APACHE, I would not recommend abandonment of a more sophisticated system that has been proven to work.

We did perform a multivariate analysis, and the results of this analysis suggest that if gone head-to-head the APACHE III score would prove a stronger system to use.

We have no fixed protocol for resuscitation in our ICU, in response to your second question, but having worked many months in that ICU I can tell you that Dr Barie does give us all special high-intensity training, and we are sort of all on the same page as far as that goes. We use heart rate monitoring, urine output, the standard monitoring as appropriate. We do have a high incidence of elderly patients and, because of this, a high incidence of coronary disease, and we do tend to use invasive monitoring, though we like many others have come away from using Swans, for obvious reasons, as often as we can. I think given the size of the database we might be able to determine a difference in our patients who were Swan vs those who weren’t, but I am not sure that this would have any clinical significance or importance.

And finally, I think in answer to your question about separating patients who were inadequately resuscitated and those in whom a proinflammatory state exists, I think it is difficult to separate the 2. It probably is a mixed population and would be very difficult to extrapolate the 2 separate populations, but I do think that the adequacy of the resuscitation in these postoperative patients may be a determinant of the balance between the anti-inflammatory response and the proinflammatory response.

Mark A. Malangoni, MD, Cleveland, Ohio: This is a fascinating study and your work is very interesting. What it brought to mind was in a lot of other scoring systems that have multiple elements, there is one element or sometimes a few elements but not all that really make the difference. Earlier in this meeting, there was discussion about the lack of sensitivity of fever. Have you looked at each of these individual components of SIRS, and if you have, can you tell us which one is the most important and are there certain elements of SIRS that perhaps we could just quit looking at?

Dr Talmor: We actually are very interested in that point ourselves. There have been several papers, especially recently, looking at heart rate, for example, as a predictor of outcome, and studies have shown that patients who are treated postoperatively with β-blockade have done better than those who aren’t. In fact we also stipulated that perhaps a combination of 2 of these criteria would portend a worst outcome, such as fever and white blood cell count, which seem to me most predictive of sepsis, and in just starting these subset analyses, which obviously involve very, very complicated statistical analyses in these 2300 patients that we have accumulated, it does seem to point to the fact that heart rate will be the most significant predictor of the 4, which is not what we would have necessarily predicted, but the analyses aren’t complete at this point.

Per-Olof Nyström, MD, PhD, Linköping, Sweden: Well, since all the SIRS criteria are contained in the APACHE II and especially APACHE III scoring, I am very happy to learn that you think we should not substitute those more sophisticated scores in terms of using them for outcome prediction; however, that begs the question, why should we use the SIRS score? What purpose, what does it tell us about patients? And I wish you could speculate a little about why you use it.

Dr Talmor: I don’t think that the SIRS score tells us more about the patient obviously than the APACHE II or the APACHE III scores, which are far more sophisticated, but as Dr Spain alluded, those of us who are calculating these APACHE scores, by hand in our situation because the software is prohibitively expensive to do this, the calculation of the scoring is extremely tedious. If we could go to a patient’s bedside, look at the bedside monitor and immediately know what the SIRS score was, and if this did give us an accurate prediction that would allow us to have a dialogue with the patient and the family regarding outcome, I think it would be a very powerful tool for us, and that is what we are hoping to do.