Validation of the Modified Multisystem Organ Failure Score as a Predictor of Mortality in Patients With AIDS-Related *Pneumocystis carinii* Pneumonia and Respiratory Failure*

David M. Forrest, MD; Ognjenka Djurdjev, Bsc; Carlos Zala, MD; Joel Singer, PhD; Lindsay Lawson, MD, FCCP; James A. Russell, MD; and Julio S.G. Montaner, MD, FCCP

**Study objectives:** To validate a previously developed multisystem organ failure (MSOF) score with and without the addition of the lactate dehydrogenase (LDH) level as a predictor of survival to hospital discharge in patients with AIDS-related *Pneumocystis carinii* pneumonia (PCP) and acute respiratory failure (ARF).

**Design:** Retrospective chart review between April 1, 1991, and September 30, 1996.

**Setting:** University-affiliated tertiary care center in downtown Vancouver, British Columbia, Canada.

**Patients:** All patients with PCP-related ARF admitted to the ICU of St. Paul’s Hospital during the study period.

**Interventions:** As putative prognostic instruments, data were extracted regarding the APACHE II (acute physiology and chronic health evaluation II), acute lung injury (ALI), AIDS, and modified MSOF scores, as well as LDH levels, at entry to the ICU. Patients were stratified based on an LDH level of < or ≥2,000 U/L and this threshold was assessed in its predictability of outcome when added to each of the above scores. For APACHE II, the score was categorized in six groups and evaluated with and without inclusion of the LDH. Receiver operating characteristic curves were constructed for LDH and for each score with and without the LDH level to assess accuracy of prediction. The area under each curve was calculated and compared to estimate the statistical significance of observed differences.

**Measurements and results:** There were 40 admissions to the ICU of 38 patients with 52.5% mortality. The ALI and AIDS scores were not predictive of outcome. The modified MSOF and APACHE II scores were significant predictors of survival and the performance of both was enhanced by the addition of LDH.

**Conclusions:** Both the APACHE II and the modified MSOF scores were significant predictors of survival in the patient population studied. These results validate the modified MSOF score as an effective predictor of survival to hospital discharge among patients with AIDS-related PCP who develop ARF and the performance of the score is enhanced by the addition of the LDH level.

(CHEST 1998; 114:199-206)

**Key words:** acquired immunodeficiency syndrome; acute respiratory failure; *Pneumocystis carinii* pneumonia; prognostic scores

**Abbreviations:** ALI=acute lung injury; APACHE=acute physiology and chronic health evaluation; ARF=acute respiratory failure; LDH=lactate dehydrogenase; MSOF=multisystem organ failure; PCP=*Pneumocystis carinii* pneumonia; ROC=receiver operating characteristic

---

*From the British Columbia Center for Excellence in HIV/AIDS (Drs. Forrest, Zala, and Montaner, and Mr. Djurdjev) and the Department of Medicine (Drs. Lawson, Russell, and Montaner), St. Paul’s Hospital; the Division of Critical Care Medicine (Drs. Forrest and Russell), Division of Respiratory Medicine (Dr. Lawson), Department of Health Care and Epidemiology (Dr. Singer), and Faculty of Medicine (Drs. Lawson, Russell, and Montaner), University of British Columbia; and the Canadian HIV Trials Network (Mr. Djurdjev and Drs. Singer and Montaner), Vancouver, British Columbia, Canada. Manuscript received July 11, 1997; revision accepted December 26, 1997. Reprint requests: Julio S.G. Montaner, MD, FCCP, BC Center for Excellence in HIV/AIDS, Room 667, St. Paul’s Hospital, 1081 Burrard St, Vancouver, BC, Canada V6Z 1Y6

*Pneumocystis carinii* pneumonia (PCP) is associated with respiratory failure in 5 to 30% of cases and has been reported to be the most common cause of acute respiratory failure (ARF) and of admission to the ICU among HIV-seropositive patients. ARF in the setting of AIDS-related PCP has been associated with a poor prognosis and high in-hospital mortality. However, this has been variable through the AIDS epidemic at least in part due to changes in patient selection for ICU support and increased use of corticosteroids as adjuvant therapy. Thus, identification of subgroups destined to have a high
mortality is important in advising patients on treatment options.

We have developed previously a prognostic scoring instrument based on a modification of a multi-system organ failure (MSOF) score and found it highly predictive of outcome of ARF secondary to AIDS-related PCP, particularly when the lactate dehydrogenase (LDH) level was added to the score. In brief, the MSOF score is calculated by assigning one point for each organ system failure. Organ systems included are respiratory, cardiovascular, renal, hepatic, and hematologic, as well as GI bleeding and encephalopathy prior to intubation. We therefore undertook the present study to reevaluate the prognostic scores assessed in our original study and to validate the modified MSOF score with and without the addition of the LDH level as a predictor of survival to hospital discharge at the time of ICU admission in a subsequent set of patients with AIDS-related PCP and ARF.

**Materials and Methods**

Assessment of the MSOF score was undertaken originally in patients in the period between January 1, 1985, and April 1, 1991. For the present study, all cases of patients with AIDS-related PCP admitted to the ICU at St. Paul's Hospital with respiratory failure between April 1, 1991, and September 30, 1996, were identified through a computerized search of hospital records. Patients with respiratory failure who were not admitted to the ICU were not included. All ICU cases were used in the analysis, whether or not mechanical ventilation was used. Repeated admissions were considered separate episodes of PCP if hospital admissions were separated by at least 30 days. Charts were reviewed by hand using a standardized case report form. As in our previous study, data were recorded regarding the APACHE II (acute physiology and chronic health evaluation II), acute lung injury (ALI), AIDS, and modified MSOF scoring instruments, as well as LDH levels, at entry to the ICU. Reviewers were not blinded as to the goals of the study. The outcome assessed was survival to hospital discharge.

Patients with both microbiologically proven PCP by bronchoscopy or at autopsy and those with presumptive PCP were included in the analysis. A presumptive diagnosis of PCP was made if HIV-seropositive patients had a compatible clinical history and physical examination, CD4+ lymphocyte count <200/mm³, hypoxemia with alveolar-arterial oxygen gradient of ≥50 mmHg, a chest radiograph showing diffuse alveolar interstitial infiltrates, and if no other diagnosis was established. ARF was defined as PaO₂/fraction of inspired oxygen <150.

Patients were stratified based on an LDH level of <2,000 U/L, to assess predictability of outcome, as previously described. This threshold was assessed in its predictability of outcome when added to the MSOF, AIDS, and ALI scores, thus increasing the range of each score by one. This threshold cannot be used for the APACHE II score, which requires gradation of degree of abnormality. Thus, as in our previous study, the threshold for LDH in this instance was set at 550 U/L, 1,650 U/L, 2,000 U/L, and 2,130 U/L, so increasing the range of the score by four points. As in our previous study, the APACHE II score was categorized in six groups.

Receiver operating characteristic (ROC) curves were constructed for each score and for LDH and then repeated for each score plus the LDH level to assess the accuracy of prediction of the scores. In brief, we plotted the true-positive rate (sensitivity) on the vertical axis of a graph against the false-positive rate (1-specificity) on the horizontal axis for each threshold. An ideal test has a high true-positive rate and a low false-positive rate and therefore should produce a curve close to the upper left corner of the graph. In contrast, a test with no predictive value should have equal numbers of true and false positive results and thus would produce a curve along the identity line along the diagonal. The area under each curve so generated was calculated and compared with curves for other scores to estimate the statistical significance of observed differences.

Categorical variables were compared using Mantel-Haenszel $\chi^2$. The MSOF score with and without modification by addition of the LDH level was further assessed as predictive of outcome (mortality) by logistic regression. All statistical tests were two sided and p<0.05 was considered statistically significant.

**Results**

Between April 1, 1991, and September 30, 1996, a total of 40 admissions to the ICU for ARF were identified in 38 patients. Complete data were available for all patients. Two patients were admitted twice to the ICU; all repeat admissions were separated by at least 30 days. A total of 37 (98%) of these patients were male and the median age was 38 years. Thirty (75%) of the cases represented a first episode of PCP and the diagnosis of PCP was confirmed microbiologically in 33 episodes (83%). Of the seven patients with presumptive PCP, four never had bronchoscopy performed and three had bronchoscopy >7 days after initiation of therapy. All patients received combined therapy with anti-PCP drugs and corticosteroids. Among the 40 patients admitted to the ICU, 25 (63%) received mechanical ventilatory support. Baseline characteristics are summarized in Table 1.

Twenty-one (52.5%) of the patients died during the index hospitalization. The mortality rate was 54.6% among patients with confirmed PCP and was 42.9% among those with presumptive PCP. This difference is not statistically significant (p=0.884).

The ALI and AIDS scores were not predictive of outcome (Table 2). As shown in Table 2, the MSOF and APACHE II scores were significantly different between those who survived compared with those who did not (p=0.012 and p=0.004, respectively). Based on a threshold of 2,000 U/L, LDH alone was not significantly different in those patients who died compared with those who survived (data not shown; p=0.31). As shown in Table 3, the predictability of mortality by the MSOF and APACHE II scores was slightly enhanced by the addition of the LDH (p=0.002 and p=0.001, respectively). The predictive value of the MSOF score (with and without LDH) was retained in the subgroup of patients who were...
mechanically ventilated, as well as those in whom the diagnosis of PCP was confirmed. Addition of the LDH to the ALI and AIDS scores enhanced their predictive accuracy as well (Table 3).

As shown in Figure 1, ROC curves were generated for each of the scores alone and for LDH. The MSOF and APACHE II scores were superior to the other scores and LDH alone, based on the larger area under the ROC curve for those plots, although there was no statistically significant difference between the areas under the ROC curves (Table 4). As shown in Figure 2, ROC curves were generated adding LDH to the MSOF, APACHE II, ALI, and AIDS scores. The predictability of the MSOF and APACHE II scores was enhanced once again by the addition of the LDH level (Table 4).

The analysis was unchanged when restricted to those patients who received mechanical ventilation and to those in whom the diagnosis of PCP was confirmed (data not shown). The MSOF and APACHE II scores remained significant predictors of mortality among patients with confirmed PCP (p=0.016 for MSOF score; p=0.009 for APACHE II), while the other scores were not. Only the MSOF and APACHE II scores were predictive of outcome for patients who received mechanical ventilation (p=0.042 for MSOF score; p=0.019 for APACHE II). The predictive accuracy of the scores was enhanced also by the addition of LDH to the scores.

A multivariate analysis was performed to assess the predictive ability of the MSOF score (with and without inclusion of the LDH) relative to the other predictive scores. With the MSOF score included in the regression analysis, no other score contributed to prediction of outcome (Table 5).
**Figure 1.** ROC curves for scoring systems without LDH and for LDH itself. Each point on the curve represents the true-positive rate (sensitivity) and false-positive rate (1-specificity) at different gradations of each score (Table 2). The curves for the MSOF and APACHE II scores are the best predictors of outcome.

**Discussion**

Our results successfully validate the modified MSOF score as a predictor of outcome for patients admitted to the ICU with AIDS-related PCP and ARF. However, in contrast to our earlier findings, we found that the APACHE II score was a significant predictor of outcome and was a slightly better predictor of outcome in the more recent patient group than the MSOF score. The LDH level alone at time of ICU admission was not predictive of outcome, but the combination of LDH with all the prognostic scores enhanced their predictive ability. Thus, addition of the LDH to the MSOF and APACHE II scores significantly improved their predictive accuracy. Similarly, the ALI score itself was not a significant predictor of outcome, but it was predictive of mortality when combined with the LDH. However, the AIDS scoring system, even in combination with the LDH level, is a poor predictor of outcome in this patient population.

Validation of prognostic scoring systems is an important component of their development. Given the intended utility of scores in AIDS-related PCP and ARF in identifying patients who are destined to have a poor outcome and thus to assist in decision-making.

**Table 4—Areas Under the ROC Curves for Scoring Systems With and Without LDH**

<table>
<thead>
<tr>
<th>Scoring System</th>
<th>Area Under ROC Curve</th>
<th>Scoring System</th>
<th>Area Under ROC Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSOF</td>
<td>0.682</td>
<td>MSOF+LDH</td>
<td>0.776</td>
</tr>
<tr>
<td>APACHE II</td>
<td>0.759</td>
<td>APACHE II+LDH</td>
<td>0.792</td>
</tr>
<tr>
<td>ALI</td>
<td>0.618</td>
<td>ALI+LDH</td>
<td>0.697</td>
</tr>
<tr>
<td>AIDS</td>
<td>0.677</td>
<td>AIDS+LDH</td>
<td>0.627</td>
</tr>
<tr>
<td>LDH</td>
<td>0.551</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*The difference between areas under the curves for different scores is not statistically significant.*
Figure 2. ROC curves for scoring systems, including LDH. The curves for the MSOF and APACHE II score + LDH are the best predictors of outcome, although addition of LDH enhances the predictability of all scores.

making in this group of critically ill patients, such validation seems especially crucial. Nevertheless, of 16 tools developed that have been proposed to identify prognostic variables in PCP, only 3 have been studied in order to validate them. This may explain, in part, the resulting reluctance of clinicians to use these scores in clinical practice. Two of the scoring systems were developed for use in patients with PCP at time of initial presentation and were not specifically designed for prognostication of patients with ARF due to PCP. While both were found predictive of mortality, neither has been assessed in a critically ill population.

The third score is the APACHE II system score of Knaus and Draper. This tool has been validated as predictive instrument in a variety of critically ill patients, but such validation was conducted prior to the rapid escalation in the AIDS epidemic. Smith and colleagues and subsequently Chu have evaluated the ability of the APACHE II score to predict outcome in patients with PCP and respiratory failure. Smith et al. assessed the system in patients with PCP and respiratory failure specifically, while Chu evaluated its performance in AIDS patients with respiratory failure due to a variety of causes. In neither case was the system found to be a good predictor of outcome, and observed mortality exceeded that predicted by the model. This may be

Table 5—Multivariate Analysis of Mortality by Logistic Regression According to MSOF Score and LDH

<table>
<thead>
<tr>
<th>Scoring System</th>
<th>β</th>
<th>Odds Ratio (95% CI)*</th>
<th>χ²</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSOF</td>
<td>1.2826</td>
<td>3.606 (1.094-11.884)</td>
<td>4.4433</td>
<td>0.0350</td>
</tr>
<tr>
<td>MSOF + LDH</td>
<td>1.4486</td>
<td>4.257 (1.435-12.630)</td>
<td>6.8168</td>
<td>0.0090</td>
</tr>
</tbody>
</table>

*CI = confidence interval.
related to the heterogeneity of the populations studied and the characteristics of the particular ICU, as well as the inaccuracy of the predictive tool in mechanically ventilated patients. A modification of the instrument by addition of LDH values has been shown to enhance the predictive accuracy of the APACHE II system, but this has not been validated. In contrast, one group found the APACHE II score to be a useful prognostic indicator in AIDS patients admitted to the hospital, but this was not specific for PCP or critically ill patients.

The APACHE II score (with or without the addition of LDH) was a modestly better predictor of outcome in our patient group. While this supports the work of Benson and colleagues, it refutes that of Smith and coworkers. Its effectiveness as a predictor in this sample and not in our previous study and inconsistency with the results of other investigators, however, suggest the sensitivity of the APACHE II score to variability in the patient population studied. It may also reflect random variation, given the relatively small sample sizes in this and the previous study and multiple comparisons, but this is unlikely, since the results are highly statistically significant (p<0.005). Similarly, the ALI score when combined with the LDH level was a significant predictor of outcome, but not when evaluated alone (p=0.054). Although this may be the consequence of relatively small sample size, it also suggests that the score is not highly or consistently predictive and that it may also be sensitive to heterogeneity of the patient population studied. This agrees with the findings of Peruzzi et al, who found that markers of oxygenation and radiographic abnormalities on admission to the ICU were not predictive of outcome. In contrast, the modified MSOF score (with or without LDH) consistently predicted outcome in two separate samples, supporting its predictive utility in the population of patients with PCP-related respiratory failure and its use as the preferred predictive physiologic score in this patient group.

Indeed, there are substantial differences between the populations studied in our original analysis and the current study. First, the era of study is different. The outcome of ARF secondary to AIDS-related

**FIGURE 3.** Comparison of distribution of scores in previous and present patient populations studied. The differences in distribution of scores (especially for the APACHE II and MSOF scores) indicate that the two patient groups are different.
PCP has changed throughout the AIDS epidemic. While ICU outcome in early years (1981 to 1985) was generally poor with survival to hospital discharge of <10%, there was substantial improvement in mortality of patients with PCP and respiratory failure between 1986 and 1988, with survival to hospital discharge of approximately 40%. However, since 1989, there has been a substantial decrease in survival to hospital discharge to ≤25%. These differences are likely due to changes in patient selection for ICU support and increased use of corticosteroids as adjuvant therapy and a dismal prognosis for patients who have failed to respond to maximal therapy. The patient set used to assess the scoring instrument initially was admitted to ICU during the period January 1, 1985, to April 1, 1991, thus spanning the three eras of ICU experience described for PCP patients with respiratory failure. In contrast, the patients included in the present study were all cared for in the modern era.

There are other clinically significant differences between the two groups. Those in the derivation study had a mortality rate of 67%, while that of the patient group in the current study was 50%. Further, all patients in the current study received combined corticosteroids and anti-PCP therapy, while only 85% of the earlier group received adjuvant steroids. Although 90% of patients in our first study required mechanical ventilation, only 63% were mechanically ventilated in the present study. Finally, there were important differences in the distribution of patients across the various score categories, particularly for the APACHE II score (Fig 3). This suggests that the modified MSOF score is resistant to secular trends in patient treatment and selection and may explain the improved performance of the APACHE II score in the present patient group. Moreover, we found that the MSOF score (with or without LDH) performed well in the subgroup of patients with PCP-related respiratory failure who required mechanical ventilation and in those in whom the episode of PCP was proven bacteriologically (data not shown). These were the only groups studied in our original sample.

The modified MSOF score reflects burden of illness in various organs. We speculate, therefore, that survival among patients with ARF secondary to AIDS-related PCP is greatly influenced by the degree of compromise and physiologic derangement of other organ systems at the time of admission to the ICU. This seems logical clinically and supports the similar predictability of physiologic scores in other critically ill populations including the performance of the APACHE II score in this study and that of Benson and colleagues. Indeed, others have found that physiologic parameters at ICU admission for similar patients, either assessing respiratory function alone or incorporating measures of severity of other organ dysfunction, are valuable predictors of outcome.

In conclusion, therefore, our results successfully validated the modified MSOF score as an effective predictor of survival to hospital discharge among patients with AIDS-related PCP in whom ARF develops. We have also confirmed that the performance of the score is enhanced by the addition of the LDH level. The consistent performance, simplicity, and clinical relevance of the modified MSOF score make it a useful and attractive tool for clinical and research purposes. Its ease of calculation should facilitate use of the instrument in prognostication for patients. This can inform a discussion about the role and desirability of pursuing critical care support given anticipated outcomes and should therefore aid patients and their caregivers in clinical decision making.

REFERENCES

14 Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143:29-36
16 Hanley J, McNeil B. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 143:839-43
30 Zaman M, White D. Serum lactate dehydrogenase levels and Pneumocystis carinii pneumonia. Am Rev Respir Dis 1988; 137:796-800