Published in final edited form as:

Shock. 2009 May ; 31(5): 438–447. doi:10.1097/SHK.0b013e31818ba4c6.

Validation of Postinjury Multiple Organ Failure scores

Angela Sauaia^{1,2}, Ernest E. Moore^{2,3}, Jeffrey L. Johnson^{2,3}, David J. Ciesla⁴, Walter L. Biffi^{2,3}, and Anirban Banerjee²

¹Department of Medicine, Division of Health Care Policy and Research, University of Colorado Denver

²Department of Surgery, University of Colorado Denver

³Department of Surgery, Denver Health Medical Center

⁴Department of Surgery, University of South Florida and Tampa General Hospital

Abstract

Most multiple organ failure (MOF) scores were developed over a decade ago, but little has been done in terms of validation and to understand the differences between populations identified by each of them. Given the lack of a gold standard, validation must rely upon objective clinical and resource utilization outcomes. Thus, we propose to: 1) validate two widely accepted MOF scores (Denver's and Marshall's) examining their association with adverse outcomes in a postinjury population; and 2) compare risk factors, characteristics and outcomes of patients identified by each score. The Denver MOF score grades (from 0-3) 4 organ dysfunctions (lung, kidney, liver, heart) and defines MOF as score > 3 . The Marshall score grades, in addition, central nervous system (CNS) and hematologic dysfunction (for a total of 6 organs on a 0 to 4 scale). Using a prospectively collected dataset, MOF scored daily by both scores for 1389 consecutive trauma patients with ISS > 15 , admitted from 1992-2004, and their outcomes evaluated (death; ventilator-free days, VFD; mechanical ventilation time, MV; and length of stay in the intensive care unit, ICU-LOS). Three major groups could be identified: 1) severe injury group for whom MOF risk factor rates, mortality and utilization were all high (Denver = Marshall = MOF and Denver = MOF + Marshall = No MOF); 2) moderate injury group with medium rate of MOF risk factors, medium utilization and low mortality (Denver = No MOF, Marshall = MOF); and 3) mild injury group for whom risk factor rates, mortality and utilization were all low (Denver = Marshall = No MOF). Both scores performed well, with the Denver MOF score showing greater specificity. The basic concepts of each score can probably be combined to produce an improved MOF score.

Keywords

adult; mortality; blood transfusion; cohort study; critical care

Introduction

Multiple organ failure (MOF) remains the major cause of late postinjury death and consumes inordinate resources in the intensive care unit (ICU).¹⁻³ Yet, statistics describing this condition vary widely. Indeed, the reported incidence of postinjury MOF varies from as low as single digit numbers to values approaching 40%, with case-fatality rates also varying within similarly large ranges.⁴⁻⁹ Studies had suggested that MOF is disappearing owing to advances in trauma and critical care while other studies have not found a consistent change in either the incidence or the mortality rate associated with postinjury MOF.¹⁰⁻¹³ These disparities are in part due to different populations observed for a relatively short time, but also to the difficulty in defining MOF. The syndrome seems to fit well with Justice Stewart's famous quote: "I may not be able to define it, but I know it when I see it."

Recently, Arthur Baue in a thorough review of the MOF literature highlighted the plethora of definitions and scoring systems and called for a common definition.¹⁴ Most definitions were developed over a decade ago, yet little has been done in terms of validating them, evaluating their potential to be used as continuous scales or, moreover, to understand how and why they differ. Indeed, Marshall, who led the construction of the multiple organ dysfunction score (MODS), noted in a recent editorial that in order to do better in measuring this condition, we need to better understand the differences between scales, and to determine whether one approach is superior, or alternatively, whether two different, but informative facets of a question are being addressed.¹⁵ Thus, a thorough validation of the scores in a homogeneous population over an extended period of time is necessary.

Since 1992, we have prospectively collected clinical data on patients at risk for postinjury MOF for the first 28 postinjury days.¹⁶⁻²² To our knowledge, this is one of the longest sustained databases on postinjury MOF using standardized data collection methods. Therefore, we propose to use this database to validate two scores (Marshall MOD Score (Marshall) and Denver MOF Score (Denver)). These two scores have been used successfully in multiple trials, which underscores the importance of validating them and understanding how their classification of patients differ, so we can better understand the body of literature using them as endpoints.²³⁻³² In the absence of a gold standard, validation will entail evaluating their association with several objective, adverse outcomes, reflecting clinical status and resource utilization. We also aimed at determining how the patients identified by each of these two scores differ in terms of risk factors, characteristics and outcomes to facilitate the interpretation of studies using different scores. Finally, we examined the scores' potential as continuous scales that could be used to monitor patients response to treatment and prognosis, a critical step in monitoring the dynamic process of organ dysfunction.

Material and Methods

Study population and data collection

acutely injured patients admitted to the Rocky Mountain Regional Trauma Center surgical intensive care unit (SICU) at Denver Health Medical Center (DHMC) were studied prospectively from 1992 through 2004. DHMC is a state designated Level I trauma center

verified by the American College of Surgeons Committee on Trauma. Inclusion criteria were Injury Severity Score (ISS) greater than 15, survival longer than 48 hours from injury, admission to the SICU within 24 hours of injury, and age greater than 15 years. Patients with isolated head injuries [head injuries with an extra-cranial abbreviated injury score (AIS) less than two] were excluded. Burn patients and hanging injuries were not included.

Patient characteristics were recorded at the time of hospital admission. Daily physiologic and laboratory data were collected through SICU day 28 and clinical events were recorded on all patients thereafter until death or hospital discharge. The data collection and storage processes are in compliance with HIPAA regulations and have been approved by our Institutional Review Board.

Multiple organ dysfunction scores

data were collected daily to calculate the Denver MOF Score (Denver score) and the Marshall MOD score (Marshall score).³³⁻³⁶ In brief, the Denver score rates the dysfunction of four organ systems (pulmonary, renal, hepatic, and cardiac), which are evaluated daily throughout the patient's intensive care unit stay and graded on a scale from 0 to 3 (Table 1), with the total score ranging from 0-12. Table 2 outlines the Marshall score, which evaluates the same four organ dysfunctions as the Denver score, using slightly different criteria, plus two additional organ dysfunctions, i.e., hematologic and neurologic. The Marshall score grades dysfunction from 0 to 4, and the total score ranges from 0-24. The values for respiratory dysfunction have been adjusted for altitude by multiplication of the value by the ratio of atmospheric pressure in Denver to that at sea level (630mmHg/ 760 mmHg). Both MOF scores were calculated as the sum of the simultaneously obtained individual organ scores on each hospital day. Daily laboratorial and physiologic values were used for calculation of both scores.

MOF status by the Denver score was defined as recommended by the authors as well as others investigators, i.e., score >3 occurring any day after 48 hours postinjury.^{37;38} The authors of the Marshall score have not recommended a specific cutoff, rather they established score ranges for evaluation of mortality, as follows: 9 to 12, 13 to 16, 17 to 20, and greater than 20 with associated mortalities of 25%, 50%, 75%, and 100%, respectively.³⁹ Other studies, including those published by the Glue Grant group (www.gluegrant.org) have used a Denver score > 3 for at least two consecutive days and a Marshall score > 5 for at least two consecutive days. In our study, we revisited all these cutoffs using ROC curves and compared the one-day and two-day duration requirement for both scores. Single organ dysfunction was defined as an organ failure grade greater than zero for both scores.

Validation of the two MOF scores and patient adverse outcomes

validation was performed by evaluating the scores' association with objective adverse outcomes occurring within 28 days postinjury. We used two hospital resource utilization outcomes_ length of ICU stay (LOS) and mechanical ventilation days (MV)_ as well as two patient specific outcomes_ ventilator-free days (VFD) and death. VFD has been recently proposed as an alternative outcome for ICU patients that accounts for patients who died

early and consequently had shorter mechanical ventilation times. It was calculated as recommended by Schoenfeld and Bernard.⁴⁰ Because LOS, MV and VFD were not normally distributed, these outcomes were categorized in LOS>14 days, MV> 7 days, and VFD<21 days, respectively. These cutoffs were chosen based on a combination of clinical relevance and the data distribution in tertiles. The prediction of these outcomes was evaluated by building logistic regression models in which the above described adverse outcomes were the dependent variables and the MOF status as defined by the scores were the independent variables or predictors. The performance of each MOF score was then evaluated using measures of goodness-of-fit of logistic regression models (1) Akaike Information Criterion, for which lower values reflect better goodness-of-fit; and 2) -2 log-likelihood reduction, for which larger reductions reflect better goodness-of-fit), sensitivity, specificity and the area under the receiver operating characteristics (ROC) curves. Finally, we evaluated both the total score as well as each individual organ dysfunction score.

Evaluation of scores as continuous scales

we evaluated the performance of each score as continuous scales by examining: 1) the correlation between mortality rates and score values at day 2 through day 7, and 2) the association of mortality with temporal changes in the score from day 2 (baseline) to days 4 and 5. The first was used to evaluate the scores as prognostic indices, while the second was used to indicate if improvement/deterioration of the score could be used to monitor patient's response to treatment. We also examined the correlation with the above mentioned score ranges for mortality evaluation proposed by the Marshall score authors.⁴¹

Defining the patients classified by each MOF score

our goal was to characterize the patients classified as: 1) MOF by both scores, 2) MOF by the Denver score (but not by the Marshall score), 3) MOF by the Marshall score (but not by the Denver score), and 4) no MOF by the two scores. For each one of these four groups we evaluated MOF risk factors, clinical outcomes (death, VFD), and utilization outcomes (MV, LOS).

Statistical analysis

all analyses were performed using SAS for Windows vs. 9.0 (SAS Institute, Cary NC) and ROC curves were constructed and the area under the ROC curve compared using StatsDirect (1990-2007 StatsDirect Limited, UK). Categorical variables were analyzed using a Chi-square test or the Fisher Exact test when expected cell values were < 5. For continuous variables with normal distribution, ANOVA or Student t-tests (with the appropriate Satterwhaite modification when the assumption of equal variances did not hold) were used. A p value < 0.05 was considered significant. Logistic regression analysis was used to adjust for multiple confounders. Data are expressed as mean, standard deviation (STD) and standard error of the mean (SEM) or, when variables were not normally distributed, as median and interquartile range (IQR, lower and upper quartiles).

Results

Overall, 1389 consecutive patients met the inclusion criteria and were prospectively followed during the 12-year study period. Table 3 depicts the characteristics of the patient population. In brief, our population consisted of mostly young, healthy males, with moderate to severe injuries, mostly due to blunt force in motor vehicle accidents. Mortality and number of VFD days were low (Table 3), but health care resources utilization was high with over one third of these patients requiring longer than 14 admission days in the surgical ICU (LOS) and/or over 7 days of mechanical ventilation.

MOF classification

Table 4 shows the risk factors distribution stratified by MOF classification, while Table 5 depicts the outcomes. As illustrated in Figure 1, three major groups could be identified: 1) a severe injury group of patients classified by both scores as MOF with high rates of risk factors, mortality and utilization, 2) a moderate injury group classified as MOF by the Marshall score but not by the Denver score with medium risk factor rates, medium utilization and low mortality, and 3) a mild injury group which included patients classified as no MOF by both scores for whom risk factor rates, utilization and mortality were all low. A fourth group of 3 to 67 patients (depending on the cutoffs and requirement for 1 or 2 days minimum duration) were classified as having MOF by the Denver MOF score but not by the Marshall score. This latter group had similar outcome rates to those observed among patients identified as having MOF by both scores (Chi square p -value >0.32 for all outcomes, except VFD <21 days, for which $p=0.054$, see Table 4 for percentages). Incidentally, logistic regression models confirmed that ISS, age and units of packed red blood cells (RBC) transfused within the first 12 hours were the major independent risk factors for MOF, regardless of the definition used. Furthermore, the magnitude of their effect was similar between the two scores [MOF defined by the Marshall score: adjusted odds ratios (OR) for ISS $>25=2.2$, OR for age >50 years $=1.5$, OR for >6 units RBC/12 hours $=4.0$; MOF defined by the Denver score: OR for ISS $>25=1.9$, OR for age >50 years $=1.8$, OR for >6 units RBC/12 hours $=3.2$]. Gender and injury mechanism were not related to MOF in this series.

Validation of the scores

the ROC curves for death, VFD < 21 days, LOS > 14 days and MV > 7 days as predicted by the Denver and Marshall MOF scores using the one-day duration requirement are shown in Figures 2 through 5. Overall, both scores were associated with areas ≥ 80 (ideal value=100). The area under the ROC curve for the Denver score was slightly greater than for the Marshall score regarding prediction of death, VFD, and mechanical ventilation longer than 7 days, although these differences were not significant, as demonstrated by the overlap in the areas' 95% confidence intervals. Using the criterion of at least two consecutive days (as opposed to > 1 day) decreased the area under the ROC curve for both the Denver and Marshall scores prediction of death (Figures 1a and 1b), as well as for the other outcomes, suggesting the two-day minimum duration criterion did not improve prediction capabilities of the scores. The ROC curves suggested that the best cutoffs for the Denver and Marshall scores were >3 and >8 , respectively. The traditional cutoff >5 for the Marshall score was associated with a worse combination of sensitivity and specificity.

Goodness-of-fit of prediction models for death was better for the Denver MOF score >3 for one day compared to the Marshall score >5 for one day ($-2\log$ likelihood reduction: 192 vs. 106; Akaike Information Criterion 610 vs. 696). Changing the Marshall score to cutoff >8 (best cutoff by the ROC curve) for death improved the goodness-of-fit of the Marshall score ($-2\log$ likelihood reduction: 162; Akaike Information Criterion: 640), but its predictive performance remained inferior to the Denver MOF score's. When the two-day minimum requirement was applied, the Denver score continued to perform better than the Marshall score, yet for both models the goodness of fit was worse than for the models using the one-day requirement ($-2\log$ likelihood reduction: 157 vs. 98.7; Akaike Information Criterion 645.1 vs. 703.3; c-statistic 0.76 vs. 0.74), confirming what the ROC curve analyses had already suggested.

Tables 6 and 7 summarize other measures of association with death and VFD <21 days for the total scores and each individual organ dysfunction, using the optimum cutoffs indicated by the ROC curves while tables 8 and 9 show the association with MV >7 days and LOS >14 days. The total scores were both associated with values of sensitivity and specificity above 70% for death and VFD <21 days but either the sensitivity or the specificity were below 70% for the MV and LOS, suggesting these scores are appropriately biased toward clinical outcomes as opposed to resource utilization.

Regarding the association of individual organ dysfunction with death, heart dysfunction, graded by either score, had the best predictive power. Kidney dysfunction, as graded by the Marshall score, also correlated well with death. This later finding suggests that the serum creatinine cutoffs used to grade kidney dysfunction in the Denver score could be adjusted to lower values, similar to those used by the Marshall score, to increase sensitivity. The association of kidney dysfunction with other adverse outcomes, however, was generally low. Not surprisingly, lung dysfunction, coded by either score, was predictive of both VFD and MV, but did not perform well with the other two outcomes. In general, hematologic and CNS dysfunctions were relatively poor predictors of all the adverse outcomes.

Use of scores as continuous scales

Figures 6a and 6b depict the correlation between the scores obtained in Days 2 (¹), 3 and 7 with mortality. The association with scores on days 4 through 6, were similar and are not shown here. The Denver score (Figure 6a) had a consistent, uniform increase in mortality risk associated with each one-point increase in the total score, particularly up to a total score of 4. Scores above four exhibited a larger variability, possibly a function of the smaller number of patients in these ranks. The Marshall's score (Figure 6b) showed a relatively small increase, in a somewhat inconsistent pattern, up to scores of 4. Above 4, the Marshall score had high variability between scores obtained in different days and lower consistency in the change in mortality risk associated with higher scores.

Using the Marshall's score proposed ranges for mortality evaluation, we observed that levels between 9-12 and 13-16 were indeed associated with mortality rates of around 25% and 50% respectively (Figure 7). Levels of 17 to 20 and above 20 were uncommon and

¹the first 24 hours were defined as day zero

associated with nearly uniform mortality. Therefore, changes of 4 points were necessary to detect an increase in the risk of death.

Improvement/deterioration of the scores over time were explored next. We examined the difference between scores obtained at Day 2 and Day 4 (Delta = Day 4 Score– Day 2 Score) and between Day 2 and Day 5 (Delta= Score at day 5 – Score at Day 2). Figure 8 illustrates the association of the scores delta with mortality. Remarkably similar numbers were observed for the delta Day 4 – Day 2 (data not shown). Patients who had increases in their total Denver MOF score of 1 to 3 points showed a consistent increase in mortality risk from 10% to about 33%. The Marshall score showed a less consistent, yet similar pattern. Of note, patients who manifested large improvements (delta = -3) were more likely to die than patients with no improvement or with smaller improvements, suggesting a relationship with the initial, baseline score. Indeed, stratifying the patients by the value of their baseline Denver MOF score (score at Day 2 $\leq$ 2 or $>$ 2) minimized this effect, as shown in Figure 9.

Discussion

The Denver MOF score was first proposed in 1991 as part of a National Institutes of Health Center Project ⁴², then modified in 1993⁴³ by a panel of expert trauma surgeons and trauma investigators. In 1994, The Marshall score was developed through a systematic review of the literature to identify which systems, and what abnormalities in those systems, defined organ dysfunction; cutoffs were derived and validated based on probability of subsequent mortality in a sample of 692 patients from a Canadian ICU. ^{36;44;45} The incidence of MOF in our study varied from 22% with a case-fatality rate of 30% according to the Denver definition, as opposed to an incidence of 50% with a case-fatality rate of 15% according to the Marshall MOF score (using the traditional $>$ 5 cutoff). Using the cutoff $>$ 8 defined as optimum by the ROC curve, the Marshall score and the Denver score had a better agreement, yet the group of patients defined as MOF by the Marshall score but not by the Denver score represented a group of patients with lower rates of mortality (6%) and VFD $<$ 21 days as well as low health services utilization. A similar pattern emerged when we looked at known risk factors for MOF, such as age, ISS and RBCs ⁴⁶⁻⁴⁹, with the group identified by both scores as having MOF having the higher rates of the risk factors, an intermediate group with medium risk factors rates who was identified as MOF by the Marshall score but not by the Denver score, and a low risk group defined by both scores as no MOF. Incidentally, once more RBC transfused in the first 6 hours emerged as major risk factor increasing the risk of MOF by at least 3 fold, regardless of definition confirming previous findings in this arena. ⁵⁰⁻⁵²

The requirement of two-day duration for the scores added complexity to both scores but did not improve the discrimination of adverse outcomes. Collectively, our findings suggest that both scores performed reasonably well, with the Denver score having a slightly better performance than the Marshall score in the discrimination of adverse clinical and utilization outcomes due to greater specificity. A similar finding was reported by Groz et al. in a recent study comparing three MOF scores (Denver, Marshall and Goris) on 301 severely injured patients who were graded daily into a group with and a group without MOF by an experienced intensive care physician.⁵³ The Denver score's sensitivity was 81% and its specificity was 88%, superior to the other scores. The authors attributed the lower specificity

of the Goris and Marshall scores to liver and cardiovascular dysfunction grading.⁵⁴ Also similar to our findings, Grotz et al. found that grading the heart, lungs, kidney and liver seemed sufficient and other organs dysfunctions did not add to the performance of the scores.

Our study found that both scores had areas under the ROC curve equal or superior to 80, a value much larger than the values recently found by Zygun et al who compared the SOFA (Sepsis-related Organ Failure Assessment) score, developed by a consensus conference at a meeting of the European Society of Intensive Care Medicine, and the Marshall score.⁵⁵ In their study, all areas under the ROC curves were below 63. A difference with our study was that trauma patients composed only 11.5% of their sample, and the majority were general ICU patients (excluding patients admitted to a general coronary unit and elective surgery with shorter stay). We found that limiting our sample to trauma patients, and further excluding isolated head injuries, burns and hanging injuries strengthened our study design by providing a homogeneous, relatively young population among whom previous medical conditions affecting major organ systems are less frequent than in the general ICU population. In addition, the time of the initial insult is often known among trauma patients, which is usually not possible in other critically ill patients.

We observed that both scores were appropriately biased toward clinical outcomes (death and VFD) as opposed to resource utilization, as measured by MV and LOS. This is not surprising as resource utilization may be influenced by factors not directly related to the patient condition, such as ICU and hospital management, ICU and hospital staff issues, protocols, availability of finite resources, etc.

In previous studies we have avoided using the Denver MOF score as a continuous measure, fearing that a scale constructed as ordinal would not serve as continuous. In this study we examined more closely the performance of the Denver MOF score as a continuous scale, which could be used to monitor patients over time, deriving prognosis and evaluating response to treatment based on temporal changes in the score. This preliminary evaluation seems to support the use of the Denver MOF score in this fashion. Denver MOF scores obtained in different days were consistently and uniformly associated with a similar mortality risk. Increases (or decreases) of a single point from day 2 to day 4 and from day 2 to day 5 were associated with measurable, consistent changes in prognosis. The Denver score seems to perform better within the range of 0 to 4, suggesting we may need to continue to refine the scale and ranges with regard to higher scores.

Overall, individual organ dysfunctions functioned inconsistently as predictors of adverse outcomes, with few exceptions. In brief, kidney dysfunction, as graded by the Marshall score, seems to correlate well with mortality, suggesting that their ranges have better discriminatory performance than the Denver MOF score. Not surprisingly, lung dysfunction scores were associated with mechanical ventilation outcomes (MV>7 days and VFD<21 days). Liver dysfunction, measured by either score, had a poor predictive performance in general, possibly because bilirubin is potentially affected by factors unconnected to the liver function (e.g., hemolysis). This finding underscores the importance of assessing other measures of liver function.

Admittedly, we did not compare grading of the cardiac dysfunction by the two versions of the Marshall score: 1) the treatment-dependent version, in which heart rate, use of inotropes and lactate levels are used to grade the dysfunction, and 2) the treatment-independent version, which uses a composite measure named pressure-adjusted heart rate (PAR= heart rate X central venous pressure (CVP) / mean arterial blood pressure).^{36;56} The authors' rationale was that this composite measure corrected cardiovascular function for physiologic support and showed the desired incremental correlation with ICU mortality rate. Values for each of the three component variables had to be measured simultaneously. According to the Marshall et al., the PAR satisfied most of the characteristics of an ideal variable, except that it is potentially altered as a result of transient changes associated with resuscitation, and may be affected by such therapeutic interventions as the use of beta-adrenergic receptor blockers or vasopressors.⁵⁷

While we agree that the treatment independence (although this independence was relative since PAR is directly affected by treatment) is an attractive feature, the component variable CVP is often not available; in fact, it was not available in half of the patients in the original authors' dataset.⁵⁸ We decided that imputation of zero values for the score (or a normal value of 8 for the CVP as more recently proposed by Marshall⁵⁹) in a large proportion of patients (due to missing CVP data) introduced more uncertainty than using the earlier version, which had a limited use of a treatment-dependent measurement (use of inotropes) as seen in Table 2. In addition, Marshall's group reverted to the 1994 simpler version of their cardiovascular dysfunction score in a subsequent study published in 2001.⁶⁰ Thus, we opted to use the 1994 version employing heart rate, inotrope use and lactate levels.⁶¹ Existing alternatives to the use of treatment-dependent variables are either missing in large proportion of patients or are under investigation.^{62;63} Interestingly, both the Marshall and Denver cardiac dysfunction scores functioned well as predictors of death with large areas under the ROC curves as well as sensitivity and specificity values above 80%.

Another limitation of our study is the inclusion of patients of a single ICU, which restricts our ability to evaluate the influence of different protocols on the scoring of MOF. The Glue grant, which employs the Denver MOF score as one of their outcome measures, and includes patients from several centers, should provide additional data to verify the applicability of our score in other trauma facilities. Finally, we use worst daily values for both scores, yet the Marshall score authors recommend the use of physiologic values measured at the same point in time every day (first morning values) to avoid capturing momentary physiologic changes unrelated to changes in the patient's underlying physiologic status. We made a decision to obtain the same values (worst values) for both scores, which may have an effect on the Marshall score.

In conclusion: 1) both scores perform reasonably well as indicators of adverse outcomes in critically ill patients, with the Denver MOF score performing slightly better due to greater specificity; 2) the Denver MOF score can be used as a continuous scale to monitor individual patient's response to treatment; and 3) the analysis of individual organ dysfunction scores suggest that concepts of the two scores can be combined to develop a superior score.

Acknowledgments

The authors gratefully acknowledge the assistance of the program manager, Ms. Victoria Bress, and the invaluable contribution of the medical staff at the Denver Health Medical Center.

Funding: This work was funded in part by National Institutes of Health/ National Institute of General Medical Sciences Center Grant Number 5 P50 GM049222 "Trauma Primes Cells".

References

1. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg*. 2005; 140:432–438. [PubMed: 15897438]
2. Sauaia A, Moore FA, Moore EE, Moser KS, Brennan R, Read RA, et al. Epidemiology of trauma deaths: a reassessment. *J Trauma*. 1995; 38:185–193. [PubMed: 7869433]
3. Teixeira PGRM, Inaba KM, Hadjizacharia PM, Brown CM, Salim AM, Rhee PM, et al. Preventable or Potentially Preventable Mortality at a Mature Trauma Center. [Article]. *Journal of Trauma-Injury Infection & Critical Care*. 2007; 63:1338–1347.
4. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg*. 2005; 140:432–438. [PubMed: 15897438]
5. Goris RJ. Prevention of ARDS and MOF by prophylactic mechanical ventilation and early fracture stabilisation. *Prog Clin Biol Res*. 1987; 236B:163–173. [PubMed: 3615468]
6. Laudi S, Donaubaue B, Busch T, Kerner T, Bercker S, Bail H, et al. Low incidence of multiple organ failure after major trauma. *Injury*. 2007; 38:1052–1058. [PubMed: 17572416]
7. Sperry JL, Frankel HL, Vanek SL, Nathens AB, Moore EE, Maier RV, et al. Early hyperglycemia predicts multiple organ failure and mortality but not infection. *J Trauma*. 2007; 63:487–493. [PubMed: 18073591]
8. Nast-Kolb D, Aufmkolk M, Rucholtz S, Obertacke U, Waydhas C. Multiple organ failure still a major cause of morbidity but not mortality in blunt multiple trauma. *J Trauma*. 2001; 51:835–841. [PubMed: 11706328]
9. Sauaia A, Moore FA, Moore EE, Norris JM, Lezotte DC, Hamman RF. Multiple organ failure can be predicted as early as 12 hours after injury. *J Trauma*. 1998; 45:291–301. [PubMed: 9715186]
10. Levine JH, Durham RM, Moran J, Baue A. Multiple organ failure: is it disappearing? *World J Surg*. 1996; 20:471–473. [PubMed: 8662137]
11. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg*. 2005; 140:432–438. [PubMed: 15897438]
12. Durham RM, Moran JJ, Mazuski JE, Shapiro MJ, Baue AE, Flint LM. Multiple organ failure in trauma patients. *J Trauma*. 2003; 55:608–616. [PubMed: 14566110]
13. Regel G, Grotz M, Weltner T, Sturm JA, Tscherne H. Pattern of organ failure following severe trauma. *World J Surg*. 1996; 20:422–429. [PubMed: 8662130]
14. Baue AE. MOF, MODS, and SIRS: what is in a name or an acronym? *Shock*. 2006; 26:438–449. [PubMed: 17047513]
15. Marshall JC. Measuring organ dysfunction in the intensive care unit: why and how? *Can J Anaesth*. 2005; 52:224–230. [PubMed: 15753490]
16. Ciesla DJ, Moore EE, Johnson JL, Cothren CC, Banerjee A, Burch JM, et al. Decreased progression of postinjury lung dysfunction to the acute respiratory distress syndrome and multiple organ failure. *Surgery*. 2006; 140:640–647. [PubMed: 17011912]
17. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. Obesity increases risk of organ failure after severe trauma. *J Am Coll Surg*. 2006; 203:539–545. [PubMed: 17000399]
18. Ciesla DJ, Moore EE, Johnson JL, Cothren CC, Banerjee A, Burch JM, et al. Decreased progression of postinjury lung dysfunction to the acute respiratory distress syndrome and multiple organ failure. *Surgery*. 2006; 140:640–648. [PubMed: 17011912]

19. Ciesla DJ, Moore EE, Johnson JL, Burch JM, Cothren CC, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg.* 2005; 140:432–438. [PubMed: 15897438]
20. Sauaia A, Moore FA, Moore EE, Norris JM, Lezotte DC, Hamman RF. Multiple organ failure can be predicted as early as 12 hours after injury. *J Trauma.* 1998; 45:291–301. [PubMed: 9715186]
21. Moore FA, Moore EE, Sauaia A. Blood transfusion. An independent risk factor for postinjury multiple organ failure. *Arch Surg.* 1997; 132:620–624. [PubMed: 9197854]
22. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma.* 1996; 40:501–510. [PubMed: 8614027]
23. Sperry JL, Frankel HL, Vanek SL, Nathens AB, Moore EE, Maier RV, et al. Early hyperglycemia predicts multiple organ failure and mortality but not infection. *J Trauma.* 2007; 63:487–493. [PubMed: 18073591]
24. Klein DJ, Derzko A, Foster D, Seely AJ, Brunet F, Romaschin AD, et al. Daily variation in endotoxin levels is associated with increased organ failure in critically ill patients. *Shock.* 2007; 28:524–529. [PubMed: 17589381]
25. Frink M, Pape HC, van Griensven M, Krettek C, Chaudry IH, Hildebrand F. Influence of sex and age on mods and cytokines after multiple injuries. *Shock.* 2007; 27:151–156. [PubMed: 17224789]
26. Nathens AB, Nester TA, Rubenfeld GD, Nirula R, Gernsheimer TB. The effects of leukoreduced blood transfusion on infection risk following injury: a randomized controlled trial. *Shock.* 2006; 26:342–347. [PubMed: 16980879]
27. Bulger EM, Jurkovich GJ, Farver CL, Klotz P, Maier RV. Oxandrolone does not improve outcome of ventilator dependent surgical patients. *Ann Surg.* 2004; 240:472–478. [PubMed: 15319718]
28. Cohn SM, Nathens AB, Moore FA, Rhee P, Puyana JC, Moore EE, et al. Tissue oxygen saturation predicts the development of organ dysfunction during traumatic shock resuscitation. *J Trauma.* 2007; 62:44–54. [PubMed: 17215732]
29. Ikossi DG, Knudson MM, Morabito DJ, Cohen MJ, Wan JJ, Khaw L, et al. Continuous muscle tissue oxygenation in critically injured patients: a prospective observational study. *J Trauma.* 2006; 61:780–788. [PubMed: 17033541]
30. Johnson JL, Moore EE, Gonzalez RJ, Fedel N, Partrick DA, Silliman CC. Alteration of the postinjury hyperinflammatory response by means of resuscitation with a red cell substitute. *J Trauma.* 2003; 54:133–139. [PubMed: 12544908]
31. Bulger EM, Jurkovich GJ, Nathens AB, Copass MK, Hanson S, Cooper C, et al. Hypertonic resuscitation of hypovolemic shock after blunt trauma: a randomized controlled trial. *Arch Surg.* 2008; 143:139–148. [PubMed: 18283138]
32. Sperry JL, Friese RS, Frankel HL, West MA, Cuschieri J, Moore EE, et al. Male gender is associated with excessive IL-6 expression following severe injury. *J Trauma.* 2008; 64:572–578. [PubMed: 18332795]
33. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma.* 1996; 40:501–510. [PubMed: 8614027]
34. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
35. Sauaia A, Moore FA, Moore EE, Lezotte DC. Early risk factors for postinjury multiple organ failure. *World J Surg.* 1996; 20:392–400. [PubMed: 8662125]
36. Marshall, JC. A scoring system for the multiple organ dysfunction syndrome (MODS). In: Reinhart, K.; Eyrich, K.; Sprung, C., editors. *Sepsis: Current Perspectives in Pathophysiology and Therapy.* Berlin: Springer-Verlag; 1994. p. 38–49.
37. Grotz M, von Griensven M, Stalp M, Kaufmann U, Hildebrand F, Pape HC. Scoring multiple organ failure after severe trauma. Comparison of the Goris, Marshall and Moore scores. *Chirurg.* 2001; 72:723–730. [PubMed: 11469095]
38. Sauaia A, Moore FA, Moore EE, Haenel JB, Read RA, Lezotte DC. Early predictors of postinjury multiple organ failure. *Arch Surg.* 1994; 129:39–45. [PubMed: 8279939]

39. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
40. Schoenfeld DA, Bernard GR. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. *Crit Care Med.* 2002; 30:1772–1777. [PubMed: 12163791]
41. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
42. Moore FA, Moore EE, Poggetti R, McAnena OJ, Peterson VM, Abernathy CM, et al. Gut bacterial translocation via the portal vein: a clinical perspective with major torso trauma. *J Trauma.* 1991; 31:629–636. [PubMed: 2030509]
43. Sauaia A, Moore FA, Moore EE, Haenel JB, Read RA. Pneumonia: cause or symptom of postinjury multiple organ failure? *Am J Surg.* 1993; 166:606–610. [PubMed: 8273837]
44. Marshall JC. Measuring organ dysfunction in the intensive care unit: why and how? *Can J Anaesth.* 2005; 52:224–230. [PubMed: 15753490]
45. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
46. Sauaia A, Moore FA, Moore EE, Norris JM, Lezotte DC, Hamman RF. Multiple organ failure can be predicted as early as 12 hours after injury. *J Trauma.* 1998; 45:291–301. [PubMed: 9715186]
47. Hensler T, Heinemann B, Sauerland S, Lefering R, Bouillon B, Andermahr J, et al. Immunologic alterations associated with high blood transfusion volume after multiple injury: effects on plasmatic cytokine and cytokine receptor concentrations. *Shock.* 2003; 20:497–502. [PubMed: 14625472]
48. Hensler T, Sauerland S, Lefering R, Nagelschmidt M, Bouillon B, Andermahr J, et al. The clinical value of procalcitonin and neopterin in predicting sepsis and organ failure after major trauma. *Shock.* 2003; 20:420–426. [PubMed: 14560105]
49. Sauaia A, Moore FA, Moore EE, Lezotte DC. Early risk factors for postinjury multiple organ failure. *World J Surg.* 1996; 20:392–400. [PubMed: 8662125]
50. Silliman CC, Moore EE, Johnson JL, Gonzalez RJ, Biffl WL. Transfusion of the injured patient: proceed with caution. *Shock.* 2004; 21:291–299. [PubMed: 15179127]
51. Aiboshi J, Moore EE, Ciesla DJ, Silliman CC. Blood transfusion and the two-insult model of post-injury multiple organ failure. *Shock.* 2001; 15:302–306. [PubMed: 11303730]
52. Moore FA, Moore EE, Sauaia A. Blood transfusion. An independent risk factor for postinjury multiple organ failure. *Arch Surg.* 1997; 132:620–624. [PubMed: 9197854]
53. Grotz M, von Griensven M, Stalp M, Kaufmann U, Hildebrand F, Pape HC. Scoring multiple organ failure after severe trauma. Comparison of the Goris, Marshall and Moore scores. *Chirurg.* 2001; 72:723–730. [PubMed: 11469095]
54. Grotz M, von Griensven M, Stalp M, Kaufmann U, Hildebrand F, Pape HC. Scoring multiple organ failure after severe trauma. Comparison of the Goris, Marshall and Moore scores. *Chirurg.* 2001; 72:723–730. [PubMed: 11469095]
55. Zygun DA, Laupland KB, Fick GH, Sandham JD, Doig CJ. Limited ability of SOFA and MOD scores to discriminate outcome: a prospective evaluation in 1,436 patients. *Can J Anaesth.* 2005; 52:302–308. [PubMed: 15753504]
56. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
57. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]
58. Marshall JC, Cook DJ, Christou NV, Bernard GR, Sprung CL, Sibbald WJ. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995; 23:1638–1652. [PubMed: 7587228]

59. Marshall JC. Measuring organ dysfunction in the intensive care unit: why and how? *Can J Anaesth*. 2005; 52:224–230. [PubMed: 15753490]
60. Cook R, Cook D, Tilley J, Lee K, Marshall J. Multiple organ dysfunction: baseline and serial component scores. *Crit Care Med*. 2001; 29:2046–2050. [PubMed: 11700393]
61. Cook R, Cook D, Tilley J, Lee K, Marshall J. Multiple organ dysfunction: baseline and serial component scores. *Crit Care Med*. 2001; 29:2046–2050. [PubMed: 11700393]
62. Wu TT, Yuan A, Chen CY, Chen WJ, Luh KT, Kuo SH, et al. Cardiac troponin I levels are a risk factor for mortality and multiple organ failure in noncardiac critically ill patients and have an additive effect to the APACHE II score in outcome prediction. *Shock*. 2004; 22:95–101. [PubMed: 15257080]
63. Edouard AR, Benoist JF, Cosson C, Mimoz O, Legrand A, Samii K. Circulating cardiac troponin I in trauma patients without cardiac contusion. *Intensive Care Med*. 1998; 24:569–573. [PubMed: 9681778]

Figure 1. Characteristics of patients classified as MOF and no MOF by the Marshall and Denver scores

Figure 2.

a: ROC curves for the Denver and Marshall's MOF scores showing prediction of postinjury mortality when criterion "score > cutoff for at least 1 day" was used

b: ROC curves for the Denver and Marshall's MOF scores showing prediction of postinjury mortality when criterion "score > cutoff for at least 2 consecutive days" was used

Figure 3. ROC curves for the Denver and Marshall's MOF scores showing prediction of postinjury ventilator-free days shorter than 21 days

Figure 4. ROC curves for the Denver and Marshall's MOF scores showing prediction of postinjury mechanical ventilation longer than 7 days

Figure 5. ROC curves for the Denver and Marshall's MOF scores showing prediction of postinjury ICU stay longer than 14 days

Figure 6.

a: Correlation between the Denver MOF score at Days 2, 3 and 7 and mortality

b: Correlation between the Marshall MOF score at Days 2, 3 and 7 and mortality

Figure 7. Marshall's scores on Days 2, 3 and 7 stratified by ranges for evaluation of mortality

Figure 8. Effect of changes in MOF scores from Day 2 to Day 5 on mortality rates

Figure 9. Mortality rates and temporal changes in the Denver MOF scores compared to baseline scores (Day 2 score), stratified by low and high baseline score

Table 1
Denver Postinjury Multiple Organ Failure Score

Dysfunction	Grade 0	Grade 1	Grade 2	Grade 3
Pulmonary PaO ₂ /FiO ₂ ratio	> 208	208 - 165	165 - 83	< 83
Renal Creatinine (umol/L)	<159	160 - 210	211 - 420	> 420
Hepatic Total Bilirubin (umol/L)	< 34	34 – 68	69 - 137	> 137
Cardiac Inotropes	No inotropes	Only one inotrope at a small dose *	Any inotrope at moderate dose or >1 agent, all at small doses *	Any inotrope at large dose or > 2 agents at moderate doses *

*
Inotrope doses (in ug/ Kg / min):

	Small	Moderate	Large
Milrinone	<0.3	0.4 -0.7	>0.7
Vasopressin	<0.03	0.03 -0.07	>0.07
Dopamine	<6	6 - 10	>10
Dobutamine	<6	6 - 10	>10
Epinephrine	<0.06	0.06 -0.15	>0.15
Norepinephrine	<0.11	0.11 -0.5	>0.5
Phenylephrine	<0.6	0.6 - 3	>3

Table 2

Marshall's Multiple Organ Dysfunction Score

Dysfunction	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4
Pulmonary PaO ₂ /F _i O ₂ ratio	>300	226–300	151–225	76–150	<75
Renal Creatinine (umol/L)	<100	101–200	201–350	351–500	>500
Hepatic Total Bilirubin (umol/L)	<20	21–60	61–120	121–240	>240
Cardiovascular	1994 version 1995 version **	Heart Rate <120	Heart Rate >140	Inotropes *	Lactate >5mmol/L
		PAR<10	PAR= 15.1-20		PAR> 30
Central nervous system Glasgow Coma Scale	15	13–14	10–12	7–9	<6
Hematologic Platelet count ×10 ³	>120	81–120	51–80	21–50	<20

* need for inotropes more than dopamine >3 ug/kg/min

** PAR: pressure-adjusted heart rate,

PAR = Heart Rate × Central Venous Pressure / mean Arterial Blood Pressure

Table 3
Characteristics of a trauma patient population admitted from 1992 to 2004

Variable	Mean	STD	SEM	Median	Lower Quartile	Upper Quartile
Demographic Characteristics						
Age (years)	37.5	16.7	0.4	34	24	46
Gender=male	1004 (72%)					
Race/Ethnicity						
White non-Latino	777(55%)					
Latino (of any race)	401(29%)					
Black non-Latino	149(11%)					
Asian	23 (2%)					
American Indian	15(1%)					
Other /unknown	24(2%)					
Comorbid conditions=Yes						
Substance abuse *	388(28%)					
Cardiovascular disease	152(39%)					
Pulmonary disease	76(20%)					
Diabetes	56(14%)					
Neurologic disorders **	47(12%)					
Immunologic disorders ***	23(6%)					
Other	14(4%)					
	114(29%)					
<i>Note: numbers will not add up to total /100% because patients may have >1 comorbidity</i>						
Injury severity variables						
Injury mechanism=blunt force	1059(76%)					
Injury Severity Score	29.5	11.3	0.3	26	22	34
RBC units/ 1 st 12 hours	4.5	8.8	0.2	0	0	6.0

Variable	Mean	STD	SEM	Median	Lower Quartile	Upper Quartile
FFP units/ 1 st 12 hours	2.3	5.2	0.1	0	0	2.0
Platelet units/ 1 st 12 hours	0.4	2.3	0.1	0	0	0
Crystalloids / 1 st 12 hours (L)	2.0	4.2	0.1	0	0	2.0
Outcomes						
Length of stay ICU (days)	15	14.6	0.4	11	5	21
Length of stay ICU>14 days	542 (39%)					
Mechanical ventilation (days)	6.2	7.4	0.2	3	0	10
Mechanical ventilation>7days	460 (33%)					
Ventilator –free days	20.4	9.3	0.2	25	15	28
Ventilator-free days<21 days	839 (60%)					
Mortality	116 (8%)					

STD: standard deviation; SEM: standard error of the mean; RBC: packed red blood cells; FFP: fresh frozen plasma; ICU: intensive care unit

* withdrawal syndrome documented in present or past admission;

** chronic seizure disorder requiring medication;

*** HIV-AIDS (as defined by the Center for Diseases Control), HIV positive, or radiation/chemotherapy/prednisone > 0.3mg/Kg/day (or equivalent dose of other steroids) in the last six months.

Table 4
Risk factors of patients identified as having or not MOF based on the Denver and Marshall MOF scores

	N	Mean ISS (SEM)	Mean Age (SEM)	Median RBC/12 hours (IQR)	Blunt trauma (%)
MOF defined as Denver >3 for 1day OR Marshall>5 for 1day *					
Denver=Marshall=MOF	304	34.5(0.8)	41.6(1.0)	4 (0-12)	75%
Denver=Marshall=No MOF	694	26.8(0.4)	36.4(0.6)	0 (0-2)	78%
Denver=MOF, Marshall=No MOF	4	32.5(7.3)	27(5.8)	0(0-1)	100%
Denver=No MOF, Marshall=MOF	387	30.5(0.5)	36.6(0.9)	2(0-7)	73%
MOF defined as Denver >3 for 1day OR Marshall>8 for 1day **					
Denver=Marshall=MOF	241	34.6(0.9)	41.6(1.1)	5(0-14)	75%
Denver=Marshall=No MOF	974	27.7(0.3)	36.4(0.5)	0(0-4)	77%
Denver=MOF, Marshall=No MOF	67	34.1(1.8)	40.4(2.2)	2(0-7)	78%
Denver=No MOF, Marshall=MOF	107	31.9(1.0)	36.9(1.7)	3(0-9)	71%
MOF defined as Denver>3 for 2 days OR Marshall>5 for 2 days ***					
Denver=Marshall=MOF	201	34.3(1.0)	42.4(1.3)	4(0-14)	76%
Denver=Marshall=No MOF	906	27.8(0.3)	36.4(0.5)	0(0-4)	77%
Denver=MOF, Marshall=No MOF	3	47.3(17)	33.3(6.1)	0(0-1)	67%
Denver=No MOF, Marshall=MOF	279	31.6(0.6)	37.8(1.0)	2(0-8)	73%
Total	1389	29.5(0.3)	37.5(0.4)	0(0-6)	76%

IQR: interquartile range;

* Marshall's traditional cutoff;

** optimum cutoff indicated by ROC curve;

*** Glue grant criteria

Table 5
Adverse outcomes of patients identified as having or not MOF based on the Denver and Marshall MOF scores

	N	Death Rate %	ICU Stay >14 days %	Mechanical ventilation >7 days %	Ventilator –free < 21 days %
MOF defined as Denver >3 for 1day OR Marshall>5 for 1day *					
Denver=Marshall=MOF	304	30	73	76	10
Denver=Marshall=No MOF	694	1	15	9	88
Denver=MOF, Marshall=No MOF	4	0	75	75	25
Denver=No MOF, Marshall=MOF	387	4	54	41	50
MOF defined as Denver >3 for 1day OR Marshall>8 for 1day **					
Denver=Marshall=MOF	241	35	72	78	9
Denver=Marshall=No MOF	974	2	26	17	79
Denver=MOF, Marshall=No MOF	67	10	78	70	18
Denver=No MOF, Marshall=MOF	107	6	63	53	37
MOF defined as Denver>3 for 2 days OR Marshall>5 for 2 days ***					
Denver=Marshall=MOF	201	35	75	83	7
Denver=Marshall=No MOF	906	3	23	15	80
Denver=MOF, Marshall=No MOF	3	33	33	67	33
Denver=No MOF, Marshall=MOF	279	7	67	54	36
Total	1389	8	39	33	60

* Marshall's traditional cutoff;

** optimum cutoff indicated by ROC curve;

*** Glue grant criteria

Table 6
Measures of the association between the Denver and Marshall MOF scores (total and individual organ dysfunction scores) and death

Individual Organ Dysfunction	Area under ROC curve, 95% CI	Optimum cutoff* (>X)	Sensitivity	Specificity
Denver MOF score	0.88	3	.79	.83
Marshall MOF score	0.86	8**	.78	.80
Denver Heart	0.88	1	.84	.88
Marshall Heart	0.85	2	.87	.82
Denver Lung	0.77	1	.86	.55
Marshall Lung	0.77	2	.80	.63
Denver Kidney	0.72	0	.51	.92
Marshall Kidney	0.80	0	.74	.79
Denver Liver	0.69	0	.60	.75
Marshall Liver	0.70	1	.49	.82
Marshall Hematologic	0.75	2	.53	.89
Marshall CNS	0.64	3	.49	.82

* optimum cutoff as indicated by ROC curve

** if traditional cutoff of >5 is used, sensitivity is 92% and specificity is 54%
 Shaded cells denote both sensitivity and specificity values > 70%.
 CI: confidence interval

Table 7

Measures of the association between the Denver and Marshall MOF scores (total and individual organ dysfunction scores) and ventilator-free days <21 days

Individual Organ Dysfunction	Area under ROC curve, 95% CI	Optimum cutoff* (>X)	Sensitivity	Specificity
Denver MOF score	0.88	2	0.71	0.88
Marshall MOF score	0.87	5	0.85	0.73
Denver Heart	0.73	0	0.92	0.53
Marshall Heart	0.81	1	0.70	0.79
Denver Lung	0.84	1	0.75	0.83
Marshall Lung	0.83	2	0.80	0.73
Denver Kidney	0.61	0	0.98	0.25
Marshall Kidney	0.66	0	0.87	0.44
Denver Liver	0.72	0	0.89	0.54
Marshall Liver	0.76	0	0.79	0.69
Marshall Hematologic	0.72	1	0.78	0.55
Marshall CNS	0.72	2	0.82	0.58

* optimum cutoff as indicated by ROC curve; CI: confidence interval
Shaded cells denote both sensitivity and specificity values > 70%.

Table 8
Measures of the association between the Denver and Marshall MOF scores (total and individual organ dysfunction scores) and MV > 7 days

Individual Organ Dysfunction	Area under ROC curve, 95% CI	Optimum cutoff* (>X)	Sensitivity	Specificity
Denver MOF score	0.87	1	.93	.64
Marshall MOF score	0.84	5	.85	.68
Denver Heart	0.71	0	.53	.88
Marshall Heart	0.78	1	.79	.65
Denver Lung	0.83	1	.86	.71
Marshall Lung	0.82	2	.76	.77
Denver Kidney	0.62	0	.27	.97
Marshall Kidney	0.65	0	.44	.84
Denver Liver	0.73	0	.57	.86
Marshall Liver	0.77	0	.73	.76
Marshall Hematologic	0.70	1	.55	.76
Marshall CNS	0.70	1	.71	.66

* optimum cutoff as indicated by ROC curve. Shaded cells denote both sensitivity and specificity values > 70%.
CI: confidence interval

Table 9
Measures of the association between the Denver and Marshall MOF scores (total and individual organ dysfunction scores) and LOS > 14 days

Individual Organ Dysfunction	Area under ROC curve	95% Confidence Interval	Optimum cutoff as indicated by ROC curve(>X)	Sensitivity *	Specificity *
Denver MOF score	0.80	0.78 to 0.82	1	.86	.64
Marshall MOF score	0.80	0.78 to 0.82	5	.80	.69
Denver Heart	0.65	0.63 to 0.68	0	.45	.86
Marshall Heart	0.77	0.75 to 0.79	1	.77	.69
Denver Lung	0.77	0.75 to 0.80	1	.76	.70
Marshall Lung	0.77	0.75 to 0.79	1	.88	.55
Denver Kidney	0.57	0.55 to 0.59	0	.20	.95
Marshall Kidney	0.60	0.58 to 0.63	0	.37	.82
Denver Liver	0.69	0.66 to 0.71	0	.50	.86
Marshall Liver	0.74	0.71 to 0.76	0	.67	.77
Marshall Hematologic	0.67	0.65 to 0.70	0	.78	.49
Marshall CNS	0.70	0.67 to 0.73	1	.70	.69

* optimum cutoff as indicated by ROC curve. Shaded cells denote both sensitivity and specificity values > 70%.

Shaded cells denote both sensitivity and specificity values > 70%.

CI: confidence interval