

# **Defining multiple organ failure after major trauma: A comparison of the Denver, Sequential Organ Failure Assessment and Marshall scoring systems**

## **Background**

Since Eiseman first used the term “Multiple Organ Failure” in 1977, many definitions and over 40 organ failure severity scoring systems have been published.<sup>1,2</sup> These scoring systems vary widely in the number and pattern of organ systems included, and in the variables used to measure dysfunction. Many of the scoring systems were developed in the 1980s and 1990s, since when a degree of convergence has been seen in score composition. Despite this, no single scoring system is accepted as the gold-standard measure for multiple organ failure (MOF). Without an accepted definition of MOF, wide variations in incidence, duration and associated mortality after trauma have been reported.<sup>3-7</sup> This makes comparisons between studies and populations difficult, and complicates evaluation of the effects of changes in management of trauma patients.

A review of recent literature on trauma outcomes showed three different MOF scoring systems are commonly used: the Denver PostInjury Multiple Organ Failure Score,<sup>8</sup> the Sequential Organ Failure Assessment (SOFA)<sup>9</sup> and the Marshall Multiple Organ Dysfunction Score (MODS).<sup>10</sup> The Denver and Marshall scores have been historically more widely used in trauma research, however, the SOFA score is in widespread use in general Intensive Care literature, and more recently has been used in a number of trauma studies.<sup>11-13</sup> The effect of the choice of score on the recorded incidence and pattern of MOF is not clear. Previous studies have compared the ability of the Denver and SOFA scores<sup>14</sup> and the Denver and Marshall scores<sup>12</sup> to predict mortality and adverse outcomes after trauma. The characteristics of the populations defined by the Denver and Marshall scores have been studied.<sup>15</sup> However,

no comparisons have been performed across the three scoring systems in a severely injured patient population.

The broad aim of this study was to establish which of the Denver, SOFA and Marshall scores was best for use to define and measure MOF in patients after severe trauma. A descriptive analysis of the patient demographics, MOF incidence, onset, duration and organ system breakdown, length of Intensive Care Unit (ICU) and hospital stay and early interventions in populations defined by each score was performed. The predictive properties of each of the scores for short-term mortality was then examined.

## **Methods**

### **Description of the MOF scoring systems**

#### ***Denver PostInjury Multiple Organ Failure Score***

The Denver score was developed by trauma experts in 1991, and included eight organ systems.<sup>16</sup> A modification in the mid-1990s reduced the number of organ systems to four – respiratory, renal, hepatic and cardiac.<sup>8</sup> The score was designed for adult trauma patients with an Injury Severity Score (ISS) > 15, who survived more than 48 hours from injury. The worst daily value for a single variable related to each organ system is scored from zero to three, with a total score greater than three denoting post-trauma MOF. This cut-off value has been demonstrated to be valid in prediction of trauma outcomes.<sup>15</sup>

#### ***Sequential Organ Failure Assessment (SOFA)***

The SOFA score was developed in 1994 using a consensus group of clinicians.<sup>9</sup> Several principles were agreed on prior to the selection of variables to measure organ dysfunction. Each variable was required to be simple, objective, easily and routinely measured in an ICU

setting, amenable to repetitive assessment, and independent of therapeutic interventions.

Additionally, variables should be continuous measures.

The final score included six organ systems – respiratory, renal, hepatic, cardiac, coagulation and central nervous system (CNS). The worst daily value for each system is scored from one to four to create an overall score. SOFA was originally developed as a descriptive score, and no cut-off value from MOF was suggested. Multiple derivative scores have been developed from the SOFA score, including ICU admission SOFA score, maximum SOFA score, the SOFA score at 48 and 96 hours after admission, and the change in scores over different time periods.<sup>17-19</sup> Analysis of general ICU populations has shown a direct correlation between increasing scores for each organ system and mortality.<sup>9</sup> The original authors have subsequently recommended a score of three or more to denote a single organ failure (SOF) and an overall total greater than five, or two or more SOFs, has been accepted and used in trauma research as a definition of MOF.<sup>11,20</sup>

### ***Marshall Multiple Organ Dysfunction Score (MODS)***

The Marshall MODS was developed in 1995 via a two-stage process. Initially, a literature review was performed to evaluate previous measures of organ dysfunction, and from this, a list of characteristic of the ideal descriptor for organ dysfunction was developed.<sup>21</sup> For the second part, a database of 700 surgical admissions was split into development and validation sets, and used to calibrate candidate measures of organ dysfunction against mortality.<sup>10</sup>

Whilst Marshall recommended the inclusion of seven organ systems – respiratory, renal, hepatic, cardiac, haematological, CNS and gastrointestinal (GI) – the GI system was not included in the final score, as no variable that met the criteria for the ideal descriptor of organ

dysfunction could be found. No single variable was found to meet the criteria for cardiac failure, and a composite measure called the pressure-adjusted heart rate (PAR) involving heart rate, mean arterial pressure, and central venous pressure was developed. A representative daily value, conventionally taken as the first morning measurement, is scores from zero to four. A cut-off value of greater than five to denote MOF has been widely used.<sup>22-24</sup>

A summary of the Denver, SOFA and Marshall scoring systems for MOF is given in Table 1.

### **Study Methods**

An observational cohort study was conducted on all adult trauma patients requiring ICU admission over an eight-year period from 2003-2011 at a single UK trauma center (John Radcliffe Hospital, Oxford). Patients who died within the first 48 hours after injury, or who were delayed tertiary transfers were excluded. Local practice physically separated the management of patients with isolated head injuries from those with multisystem trauma. Patients with isolated head injuries who required ICU admission for management of that injury were admitted to a separate Neurosciences ICU. Patients with multiple trauma that required ICU admission for other reasons were admitted to the Adult ICU. Isolated head injuries were therefore excluded by removal of patients treated in the Neurosciences ICU.

A new trauma database was created, involving merging of multiple data sources. Three main electronic sources were used – the national databases of the Trauma Audit Research Network (TARN) and the Intensive Care National Audit and Research Centre (ICNARC), and a system linked to electronic ICU monitoring, called IntelliVue Clinical Information Portfolio (ICIP, Phillips).

The TARN database contains details of all UK trauma admissions, and is focused on injury and trauma management data. All trauma patients regardless of age are entered into TARN, providing criteria for level of medical care and injury severity are met. The criteria for injury severity are not based on ISS, but on a descriptive list of injury parameters as given in the TARN Procedures Manual,<sup>25</sup> which equate to a moderate or severe level of injury. Meeting the TARN inclusion criteria was therefore used as the injury inclusion criteria for the study, although data on ISS were also collected. The ICNARC database is focused on patient demographics, physiological details within the first 24 hours of ICU admission, and outcome measures. ICIP is an electronic patient clinical record system which is linked to individual patient ICU monitoring, and provides extensive details on demographics, laboratory results, investigations, pharmacy, interventions, clinical and nursing notes, and hourly monitoring of physiological parameters. Permission was obtained from the local Research Ethics Committee (11/SC/0015) and the National Information Governance Board (3-04(X)/2011) for the use of data without patient consent.

Study data for the database were collected on patient demographics, injury details, timings of the clinical pathway, laboratory results, hourly physiological data, and medical and surgical interventions. The primary outcome measure was mortality within the first 100 days from injury. Additional data were taken from local electronic sources from radiology, ICU and trauma, and from written hospital notes, to minimise any missing data, and remove potential bias from data not missing at random. The study sized was based on calculations for use of the database to build predictive models from MOF, where a minimum of 420 patients was required for a 20-variable model. All patients meeting the study inclusion criteria were included in the trauma database.

In keeping with the principles of the Denver score, MOF could only be diagnosed at more than 48 hours from injury, to avoid physiological derangement resulting from inadequate resuscitation. Data on organ dysfunction variables required for calculation of the Denver, SOFA and Marshall scores were collected from day three to day 100 post-injury or until ICU discharge. Scores were calculated daily for each component organ system score, and summed for a daily overall organ failure score. Original data collection rules were adhered to, using the worst daily values for the Denver and SOFA scores, and the first morning value for the Marshall score. The previous day's value was used to substitute for any residual missing data for all three scores. MOF was defined as a Denver score  $> 3$ , a SOFA score  $> 5$ , or a Marshall score  $> 5$ . Patients were defined as having developed MOF if threshold scores were obtained for one or more days. Each patient could therefore be categorised as either having MOF or not having MOF by each of the Denver, SOFA and Marshall scores. Additional data were collected on the timing of MOF onset, duration and the individual organ systems affected.

Statistical analysis was performed using SPSS Statistics Version 23.0 with additional testing via VassarStats. For the descriptive analyses of the MOF groups, continuous variables were compared using t-tests, and categorical data using chi-squared testing. Non-normally distributed variables were compared using the Mann-Witney U test. Kaplan-Meier survival curves were used to compare mortality to day 100 post-injury in MOF and no MOF groups. Sensitivity analyses and receiver operating characteristic (ROC) curves were used to assess the predictive value of a MOF diagnosis for trauma mortality. The Area Under the Receiver Operating Characteristic (AUROC) Curve was calculated in SPSS using the non-parametric method.

## **Results**

## **Study population**

In total, 491 patients were suitable for inclusion in the trauma database from 2003-2011, all of whom were included in the MOF analyses. Patients were predominantly male (75.6%) with a mean age of 38.0 years. Road traffic collisions were the predominant mechanism of injury (68.5%). Injury severity was high, with a mean ISS of 25.3 and a mean new Injury Severity Score (nISS) of 34.7. Significant injuries with an Abbreviated Injury Score (AIS) of two or more were common to the head (59.6%), chest (59.6%) and pelvis or limbs (63.7%). Mean ICU stay was 7.6 days, and mean hospital stay was 29.8 days. Further population demographics are given in Table 2.

## **MOF group comparisons**

Overall, 202 (41.1%) of the patients did not develop MOF by any of the three scoring systems. MOF incidence was 22.8% (112 patients) with the Denver score, 40.5% (199 patients) with the Marshall score, and 58.5% (287 patients) with the SOFA score. A core group of 101 (20.6%) patients were diagnosed with MOF by all three scoring systems. A diagrammatic representation of the relationship between groups defined by the different MOF scores is given in Figure 1.

The overall patient demographics with comparison of MOF and no MOF groups by the different scoring system definitions are given in Table 2. MOF defined by the SOFA and Marshall scores was associated with increasing age, but no association was found between age and Denver MOF. All definitions of MOF were associated with a higher level of injury severity, and with increased ICU and hospital stays. SOFA and Marshall MOF were associated with the need for operative management, whilst Denver MOF was associated with the need for early transfusion.

A delay to ICU admission, as defined by admission on or after Day 2 post-injury, was associated with an increased incidence of MOF regardless of definition. Over 20% patients developing MOF by each of the three scores were admitted on or after day two following their injury. From the 83 late admissions, MOF incidence was 36.1% by the Denver score, 73.5% by the SOFA score and 50.6% by the Marshall score.

SOFA MOF was characterised by an early onset, and was present in 79.8% of the patients who developed it by Day 3 post-injury. In comparison, only 45.5% of the patients developing Denver MOF met the criteria at the same time after injury. No significant differences were observed in the duration of MOF across the three definitions. The breakdown of single organ failures showed higher numbers of patients with respiratory and hepatic failure with the Denver and SOFA scores, and higher numbers with cardiac failure with the Marshall score. (Table 3)

### **MOF and mortality**

There were 52 deaths (10.6%) within the first 100 days from injury. Significantly lower survival rates were seen in patients with MOF across all definitions, with the lowest absolute survival rate in those patients diagnosed with MOF by the Denver score. (Figures 2:A-C; Table 4) A diagnosis of Denver MOF was associated with the highest increased risk of death (Hazard Ratio 3.87; 95% CI 2.24-6.66,  $p < 0.001$ ), with smaller but significant increases in risk of death seen with SOFA MOF (Hazard Ratio 1.40, 95% CI 1.03-1.91,  $p = 0.030$ ) and Marshall MOF (Hazard Ratio 1.69; 95% CI 1.27-2.25,  $p < 0.001$ ). No deaths were seen after day 20 post-injury in patients who did not develop MOF.

Assessment of the different MOF scores as mortality predictors showed similar values for the Area Under the Receiver Operating Characteristic curve (AUROC). However, differences were observed in the relative predictors, with the Denver score showing the highest specificity and the SOFA score the highest sensitivity for post-trauma mortality. The Denver MOF score showed the highest positive predictive value for mortality, although values were low for all scores in this category. (Table 4)

## **Discussion**

Difficulties in defining post-trauma MOF have led to the creation of a large number of scoring systems. Since MOF represents the severe end of a spectrum of disease processes, many of which are inter-related, it is unsurprising that whilst some areas of consensus have developed in MOF scoring, no single definition has been accepted as a gold-standard. The majority of MOF scoring systems evolved as descriptive measures, although many have subsequently been validated and used as outcome predictors. However, few direct comparisons of different MOF definitions have been studied, making it hard to know whether similar trauma populations are being described.

The incidence of MOF in this UK trauma population was similar to that seen in other international studies. A US study in 2009 recorded a Denver MOF incidence of 22.2% and a Marshall MOF incidence of 49.7%.<sup>15</sup> A more recent Australian study recorded a Denver incidence of 15.0% and a SOFA MOF incidence of 62.7%.<sup>14</sup> It is clear that the definition of MOF has a marked impact on incidence, and therefore comparisons of populations using different MOF scoring systems are problematic.

Sauaia et al (2009) postulated that a core group of “high-risk” patients were positive for both Denver and Marshall MOF.<sup>15</sup> In this study cohort, 101 patients (20.6%) were defined as having post-trauma MOF by all three scoring systems. Only 11 patients (2.2%) with Denver MOF fell outside of this groups, indicating that the Denver score may help to more accurately identify the high-risk patient.

Breakdown of the composition of organ dysfunction seen in the study cohort showed differences depending on the score used. The major difference between the SOFA and Marshall scores, and the Denver score, is in the inclusion of coagulopathy and neurological dysfunction. Both of these organ systems have potential complications with accurate assessment of dysfunction after trauma. Coagulopathy will be impacted by whether the patient was affected by Acute Coagulopathy of Trauma Shock (ACoTS), and the extent and composition of subsequent haematological resuscitations. Neurological dysfunction may be affected by direct trauma injury, and can be difficult to accurately assess in a sedated and ventilated patient. Of note, for three of the remaining organ systems, the same variables were used to defined dysfunction, albeit with different cut-off values for scoring –  $\text{FiO}_2/\text{PaO}_2$  for respiratory failure, creatinine levels for renal failure and bilirubin levels for hepatic failure. The major difference between the scores is in the definition of cardiac failure – with the Denver score using treatment parameters, the Marshall score using physiological parameters, and the SOFA score a mixture of the two. However, it can be clearly observed that these distinctions, both in cut-off levels, and in the definition of cardiac failure, result in differing levels of MOF incidence.

It is evident that MOF, however it is defined, remains a significant problem in the severely injured patient.<sup>7</sup> Our data showed a strong association between MOF of all definitions and

both patient outcomes (mortality), and resource allocation (ICU and hospital length of stay). It is notable that similar AUROC values were obtained for all three scoring systems, despite marked differences in predictive parameters. The increased specificity of the Denver score allows for accurate prophylactic targeting of high value ICU resources, however, the lower sensitivity increases the risk of potentially missing affected patients. Denver MOF had the closest association with post-trauma mortality, and overall, it can be seen that the Denver score would appear to target a severely vulnerable and high-risk group of patients.

Admission to ICU on or after Day 2 post-injury was strongly associated with MOF development, regardless of definition. The reasons for this cannot be established from our retrospective study. Whilst the number of patients affected in this way was small (5-10 patients/year), the incidence of MOF was markedly higher in this group. This represents an important group for potential modification of their clinical management. Comparison to other trauma studies has shown that the timing of ICU admission is rarely recorded, and therefore the standard of practice in this regard is unclear.

There are several limitations to this study, since it was based on a cohort from a single trauma centre. Data completeness was dependent on electronic data sources, and although every attempt was made to identify missing data from other sources, this was not always possible. Overall mortality in the cohort was low, and small changes in the distribution of deaths may impact the relationship between a diagnosis of MOF and post-trauma mortality. The study population were admitted over an 8-year period, and trauma care may have evolved during this period. However, we know of no significant step changes in practice at our institution during this time. Whilst data from the Inflammation and Host Response to Injury Collaborative Program showed a decrease in Denver MOF incidence over this time period,<sup>7</sup>

analysis of the database from our institution did not reflect this, and no temporal changes in incidence were noted across the study period for any of the three scoring systems. We appreciate that with smaller population numbers, temporal trends are harder to identify, and accept this as a potential limitation. The study population reflects the pattern of trauma prevalent in the UK, being predominantly blunt trauma from road traffic collisions, and therefore the results may not be generalizable to populations with high levels of penetrating trauma.

Despite the limitations of the study, and the limitations of any single MOF scoring system, we would strongly advocate standardisation of a post-trauma MOF definition as crucial to allow comparison between trauma studies and populations, and to facilitate risk-adjustment and stratification. Based on the evidence from this study, we would recommend use of the Denver PostInjury Multiple Organ Failure Score as the most appropriate measure of MOF in trauma patients for a number of reasons. First, the Denver score is the simplest to calculate, making it easy to use at the patient's bedside. Secondly, the absence of measures of coagulopathy and neurological dysfunction removes the issues relating to measurement of these organ systems in the trauma patient. Thirdly, it appears to identify a particularly high-risk group of patients, which represent a subset of those identified by the SOFA and Marshall scores. Fourthly, it demonstrates high specificity for mortality, allowing for efficient resource allocation. Finally, it shows the strongest association with post-trauma mortality, with a positive Denver MOF diagnosis equating to a more than threefold increased risk of death.

## **Conclusion**

Post-trauma MOF continues to have a significant impact on patient outcomes and resource allocation. A standardised definition of MOF after major trauma is necessary to allow

accurate comparison of studies and populations, since the Denver, SOFA and Marshall scores showed differences in MOF incidence, organ dysfunction patterns, duration and mortality prediction. Whilst the nature of MOF means that no single definition is without problems, we would advocate use of the Denver PostInjury Multiple Organ Failure Score as a gold-standard measure of post-trauma MOF.

### **Author Contributions**

All authors were involved in the study design. LH conducted the original literature review, created the trauma database, analysed and interpreted the data with supervision from PW, DY and KW. All authors contributed to the writing and revision of the manuscript.

### **Acknowledgements**

LH was supported by the NIHR Biomedical Research Centre, Oxford, in her role as a trauma research fellow.

### **References**

1. Eiseman B, Beart R, Norton L. Multiple organ failure. *Surg Gynecol Obstet.* 1977;144(3):323-326.
2. Baue A. MOF, MODS and SIRS: what is in a name or acronym? *Shock.* 2006;26(5):438-449.
3. 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 Acute Care Surg.* 2001;51(5):835-841.
4. Goris R, Boekhorst T, Nuytinck J, Gimbriere J. Multiple organ failure. Generalized autodestructive inflammation? *Arch Surg.* 1985;120(10):1109-1115.

5. Ciesla D, Moore EE, Johnson J, Burch J, Cothren C, Sauaia A. A 12-year prospective study of postinjury multiple organ failure: has anything changed? *Arch Surg.* 2005;140(5):432-438.
6. Regel G, Grotz M, Weltner T, Strum JA, Tscherne H. Pattern of organ failure following severe trauma. *World J Surg.* 1996;20(4):422-429.
7. Sauaia A, Moore EE, Johnson J, Chin T, Banerjee A, Sperry J, Maier R, Burlew CC. Temporal trends of postinjury multiple organ failure: Still resource intensive, morbid and lethal. *J Trauma Acute Care Surg.* 2014;76(3):582-593.
8. Moore FA, Sauaia A, Moore EE, Haenel JB, Burch JM, Lezotte DC. Postinjury multiple organ failure: a bimodal phenomenon. *J Trauma Acute Care Surg.* 1996;40(4):501-510.
9. Vincent J, Moreno R, Takala J, Willatts S, de Medonca A, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Med.* 1996;22(7):707-710.
10. Marshall J, Cook D, Christou N, Bernard G, Sprung C, Sibbald W. Multiple organ dysfunction score: a reliable descriptor of a complex clinical outcome. *Crit Care Med.* 1995;23(10):1638-1652.
11. Brattstrom O, Granath F, Rossi P, Oldner A. Early predictors of morbidity and mortality in trauma patients treated in the intensive care unit. *Acta Anaesth Scand.* 2010;54(8):1007-1017.
12. Borgman MA, Spinella PC, Holcomb JB, Blackburne LH, Wade CE, Lefering R, Bouillon B, Maegele M. The effect of FFP:RBC ratio on morbidity and mortality in trauma patients based on transfusion prediction score. *Vox Sang.* 2011;101(1):44-54.

13. Maegele M, Lefering R, Yucel N, Tjardes T, Rixen D, Paffrath T, Simanski C, Neugebauer E, Bouillon B. Early coagulopathy in multiple injury: An analysis from the German Trauma Registry on 8724 patients. *Injury*. 2007;38(3):298-304.
14. Dewar D, White A, Attia J, Tarrant S, King K. Comparison of postinjury multiple-organ failure scoring systems: Denver versus Sequential Organ Failure Assessment. *J Trauma Acute Care Surg*, 2014;77(4):624-629.
15. Sauaia A, Moore EE, Johnson J, Ciesla D, Biffl W, Banerjee A. Validation of postinjury multiple organ failure scores. *Shock*. 2009;31(5):438-447.
16. Moore F, Moore EE, Poggetti R, McAnena OJ, Peterson VM, Abernathy CM, Parsons PE. Gut bacterial translocation via the portal vein: a clinical perspective with major torso trauma. *J Trauma*. 1991;31(5):629-636.
17. Vincent JL, de Mendonca, Cantraine F, Moreno R, Takala J, Suter PM, Sprung CL, Colardyn F, Blecher S. Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. *Crit Care Med*. 1998;26(11):1793-1800.
18. Moreno R, Vincent JL, Matos R, de Mendonca A, Cantraine F, Thijs L, Takala J, Sprung C, Antonelli M, Bruining H, Willatts S. The use of maximum SOFA score to quantify organ dysfunction/failure in intensive care: results of a prospective, multicenter study. *Intensive Care Med*. 1999;25(7):686-696.
19. Ferreira F, Bota D, Bross A, Melot C, Vincent J. Serial evaluation of the SOFA score to predict outcome in critically ill patients. *JAMA*. 2001;286(14):1754-1758.
20. Ulvik A, Kvale R, Wentzel-Larsen T, Flaatten H. Multiple organ failure after trauma affects even long-term survival and functional status. *Crit Care*. 2007;11(5):R95.
21. Marshall J. Multiple Organ Dysfunction Syndrome (MODS). In: Sibbald WJ, Vincent JL. Eds. *Clinical Trials for the Treatment of Sepsis*. Berlin. Springer-Verlag, 1995.

22. Watson GA, Sperry JL, Rosengart MR, Minei JP, Harbrecht BG, Moore EE, Cuschieri J, Maier RV, Billiar TR, Peitzman AB; Inflammation and Host Response to Injury Investigators. Fresh frozen plasma is independently associated with a higher risk of multiple organ failure and acute respiratory distress syndrome. *J Trauma*. 2009;67(2):221-227.
23. Beilman G, Blondet J, Nelson T, Nathens A, Moore F, Rhee P, Puyana J, Moore EE, Cohn S. Early hypothermia in severely injured trauma patients is a significant risk factor for multiple organ dysfunction syndrome but not mortality. *Ann Surg*. 2009;249(5):845-850.
24. Winfield R, Delano M, Lottenberg L, Cendan J, Moldawer L, Maier R, Cuschieri J. Traditional resuscitative practices fail to resolve metabolic acidosis in morbidly obese patients after severe blunt trauma. *J Trauma*. 2010;68(2):317-330.
25. The Trauma Audit and Research Network Procedures Manual: *The Healthcare Commission and The University of Manchester*.  
<https://www.tarn.ac.uk/content/downloads/53/Procedures%20manual.pdf>