Value of anatomic and physiologic scoring systems in outcome prediction of trauma patients

Mohamed Kahloul[a,c], Wahid Bouida[b,c], Hamdi Boubaker[b,c], Semir Toumi[a,c], Mohamed H. Grissa[b,c], Amira Jaafar[c], Moez Louzi[a,c], Riadh Boukef[b,c], Mourad Gahbiche[a,c] and Semir Nouira[b,c]

Objective **The goal of this study was to compare the prediction performance of two anatomic scales, the Injury Severity Scale (ISS) and the New Injury Severity Scale (NISS), with two physiologic scales, the Revised Trauma Scale (RTS) and the Simplified Acute Physiology Scale II (SAPS II), in trauma patients.**

Design **Prospective study carried out over a 16-month period.**

Setting **Emergency department of a teaching hospital.**

Patients **Hospitalized victims of trauma up to 14 years of age.**

Interventions **The primary endpoint was the survival status at hospital discharge; the secondary outcome was need for ICU admission. Model discrimination was evaluated by the area under the receiver-operating characteristic curve and model calibration was evaluated using the Hosmer–Lemeshow goodness-of-fit statistic.**

Measurements and main results **A total of 1136 patients, with an average age of 37.6 years, fulfilled the inclusion criteria. The mortality rate was 4.5%. The combined rate of hospital death and ICU admission was 17.3%. The ISS and the NISS showed excellent discriminative power for mortality prediction (AUC 0.94 and 0.93, respectively) and ICU admission decision (0.91 and 0.89, respectively), and a good calibration. The SAPS II and the RTS showed lower discriminative power. Combining ISS or NISS with SAPS II did not improve significantly the predictive performance of each scale alone.**

Conclusion **Both ISS and NISS showed better predictive severity performance compared with RTS and SAPS II in trauma patients. The combination of anatomic scales with physiologic ones did not improve the prediction performance of each scale considered alone.** ***European Journal of Emergency Medicine*** **00:000–000 © 2013 Wolters Kluwer Health | Lippincott Williams & Wilkins.**

European Journal of Emergency Medicine 2013, **00**:000–000

Keywords: emergency department, Injury Severity Scale, score, trauma

[a]Department of Anesthesiology and Surgical Intensive Care Unit, [b]Emergency Department and Research Unit UR06SP21, Fattouma Bourguiba University Hospital and [c]Department of Medicine, Faculty of Medicine, University of Monastir, Monastir, Tunisia

Correspondent to Semir Nouira, MD, Emergency Department and Research Unit UR06SP21, Fattouma Bourguiba University Hospital, Monastir 5000, Tunisia
Tel: +216 986 77343; fax: +216 73 460 678; e-mail: semir.nouira@rns.tn

Received 3 January 2013 **Accepted** 21 March 2013

Introduction

Trauma is an ever-increasing health problem worldwide. It is a major cause of mortality in the general population and especially in young individuals [1–5]. The reported mortality rates of severely injured patients range from 7 to 45% [5–9]. Many severity scales have been published to classify trauma patients in emergency or ICUs. The utility of these scores is far greater for research and surveillance purposes than their clinical utility. One of the main purposes of scoring systems in trauma is to be able to describe the trauma population objectively so that outcomes can be compared across centers and countries. Their plethora indicates not only the need for such instruments but also their shortcoming in fulfilling all requirements [6]. Furthermore, most of them have been developed and validated in countries that have their own epidemiological and demographic specificities. Consequently, before using any scoring system in different settings, its validity has to be verified.

This prospective study tested the validity of four trauma score systems: two anatomic scales, the Injury Severity Scale (ISS) [10] and the New Injury Severity Scale (NISS) [11], with two physiologic scales, the Revised Trauma Scale (RTS) [12,13] and the Simplified Acute Physiology Scale II (SAPS II) [14].

Patients and methods

Prospective data were collected on all trauma patients consecutively admitted to the emergency department (ED) of Monastir University Hospital (Tunisia) over a 16-month period (November 2008 to February 2010). We enrolled all patients presenting to the ED with blunt or penetrating trauma injuries. We did not include patients younger than 14 years of age, patients dead on arrival to the ED, patients injured as a result of burns, and those referred to another hospital. For each patient, we collected demographic data, injury mechanism, preinjury American Society of Anesthesia physical status classification [15], data required for scores calculation, intensive care admission need, length of hospital stay, and survival status at hospital discharge. ICU admission was decided at the discretion of the treating physician. It was defined as direct admission to the ICU from the ED or intensive care less than 24 h after hospital admission following surgery. This study was approved by the Ethics Committee of our institution, which waived patients' written consent because the study did not interfere with the patient's management. Baseline data and clinical course were recorded at ED admission using a uniform data collection sheet. All the data were prospectively collected by an emergency physician or by a supervised resident in the emergency medicine training program.

The Abbreviated Injury Scale (AIS) [10,11] was used to define the severity of separate injuries. To compute the ISS, each of six anatomical regions is scored with the highest AIS using the square of the highest value of the three most severely injured body regions. The total score ranged from 1 to 75 [10]. The NISS is the sum of the squares of the three highest AIS values irrespective of the anatomic area, with a total score ranging from 1 to 75. The RTS included three variables: the Glasgow Coma Scale, respiratory rate, and systolic blood pressure. A value from 0 to 4 is assigned for each variable. From these three coded values, a score is generated, ranging from 0 to 12, with lower scores representing increasing severity [12,13]. Calculation of SAPS II is based on 12 routine physiologic measurements recorded during the first 24 h after admission, previous health status, and ongoing surgical status. It resulted in an integer point score between 0 and 163 [14]. When feasible, some patients were assessed independently by a second emergency physician to judge interobserver agreement.

Statistical analysis

Continuous variables were reported as mean ± SD (or 95% confidence interval) or median and range when appropriate. Categorical data were expressed as frequency distributions. Model discrimination was evaluated by the area under the receiver-operating characteristic (ROC) curve and model calibration was evaluated using the Hosmer–Lemeshow goodness-of-fit statistic. The ROC statistic is a general measure of the test's power to separate two mutually exclusive populations, in this study, survivors and nonsurvivors. A value of 1 corresponds to a model or test that perfectly separates two populations. Values greater than 0.90 represent high accuracy of discriminating between survivors and nonsurvivors, and a range of 0.70–0.89 represents moderate accuracy. A value of 0.50 indicates that the test has no better discriminatory ability than chance. Comparison between areas under ROC curves was carried out using the z statistic according to methods of Hanley and McNeil. Calibration of the models was assessed by Lemeshow and Hosmer's goodness-of-fit statistics. The observed and expected number of deaths were compared and evaluated statistically by formal goodness-of-fit testing to determine whether the discrepancy was acceptably small ($P > 0.01$). All statistical analyses were carried out using SPSS version 13.0 (SPSS Inc., Chicago, Illinois, USA). A P value of less than 0.05 was considered significant.

Results

During the study period, 1136 patients fulfilled the inclusion criteria. Patients' demographics including age, sex, previous health status, trauma characteristics, and severity indices values are listed in Table 1. Trauma mechanisms were essentially traffic accident (27.7%), fall (24.8%), work-related accident (18.1%), domestic accident (16.3%), and violence (12.9%). Only 5.1% of patients were admitted to the ED through the emergency service facility. In this study, 191 (16.8%) patients required ICU admission with a median length of stay of 12.4 days (range, 1–63 days). The median length of hospital stay was 7.4 days (range, 1–116 days). Hospital mortality rate was 4.5% ($n = 51$). Death was directly related to the severity of injury in 24 cases. The other causes of death were septic complications ($n = 22$), iatrogenic complications ($n = 3$), and pulmonary embolism ($n = 2$). The combined death and ICU admission rate was 17.3%. Patients' severity as assessed by the different scales is summarized in Tables 2 and 3. With respect to the mortality rate, ISS and NISS achieved the highest values of area under ROC curves (0.94 and 0.93, respectively), which were superior to those of SAPS II and RTS (0.84 and 0.70, respectively, $P < 0.05$) (Fig. 1). The ISS showed the best calibration for combined events ($\chi^2 = 3.53$, $P = 0.47$) (Table 2).The same trend of results was found when these scales were assessed on their ability to predict ICU admission, with area under ROC curve values of 0.91 and 0.89, respectively, for ISS and NISS versus 0.73 and 0.58, respectively, for SAPS II and RTS (Table 3 and Fig. 2). The combination of anatomic

Table 1 Patients' characteristics

Characteristics	Value
Age [means (SD)] (years)	37.6 (20.4)
Sex	
Male [*n* (%)]	901 (79.3)
ASA physical status	
ASA I	936 (82.4)
ASA II	119 (10.5)
ASA ≥ III	81 (7.1)
Mechanism of trauma [*n* (%)]	
Traffic accident	315 (27.7)
Fall	282 (24.8)
Domestic accident	186 (16.3)
Work-related accident	206 (18.1)
Violence	147 (12.9)
Length of stay [median (range)] (days)	7.4 (1–116)
ICU admission [*n* (%)]	191 (16.8)
Mortality [*n* (%)]	51 (4.5)

ASA, American Society of Anesthesia.

Table 2 **Performance of trauma severity indices in the prediction of hospital death**

	Median (range)	Area under ROC curve (95% CI)	Hosmer–Lemeshow χ^2	Hosmer–Lemeshow *P*
ISS	9 (1–75)	0.94 (0.92–0.96)	3.43	0.48
NISS	9 (1–75)	0.93 (0.90–0.96)	4.57	0.47
SAPS II	11 (8–73)	0.84 (0.79–0.91)	11.66	0.02
RTS	11 (4–12)	0.70 (0.61–0.79)	48.16	<0.01
ISS + SAPS II	19 (9–148)	0.95 (0.93–0.97)	9.73	0.20
NISS + SAPS II	21 (9–148)	0.95 (0.92–0.98)	7.06	0.42

CI, confidence interval; ISS, Injury Severity Scale; NISS, New Injury Severity Scale; ROC, receiver-operating characteristic; RTS, Revised Trauma Scale; SAPS II, Simplified Acute Physiology Scale II.

Table 3 **Performance of trauma severity indices in the prediction of intensive care unit admission**

	Median (range)	Area under ROC curve (95% CI)	Hosmer lemshow χ^2	Hosmer lemshow *P*
ISS	9 (1–75)	0.91 (0.89–0.93)	3.53	0.47
NISS	9 (1–75)	0.89 (0.87–0.91)	31.26	<0.01
SAPS II	11 (8–73)	0.73 (0.69–0.77)	42.32	<0.01
RTS	11 (4–12)	0.58 (0.53–0.63)	603.17	<0.01
ISS + SAPS II	19 (9–148)	0.88 (0.86–0.90)	31.80	<0.01
NISS + SAPS II	21 (9–148)	0.87 (0.85–0.90)	10.64	0.15

CI, confidence interval; ISS, Injury Severity Scale; NISS, New Injury Severity Scale; ROC, receiver-operating characteristic; RTS, Revised Trauma Scale; SAPS II, Simplified Acute Physiology Scale II.

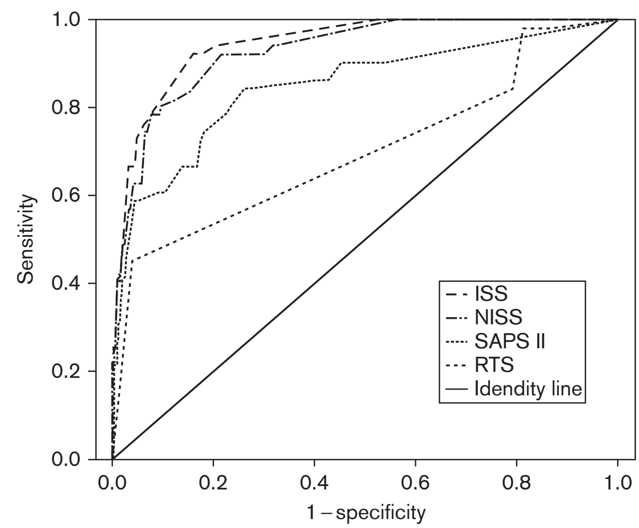
scales with SAPS II did not improve the performance of each scale considered separately (Tables 2 and 3).

Discussion

Trauma is currently one of the most major health problems worldwide. It is the disease of the young and the leading cause of death up to the age of 45 years [5–7]. In this prospective study, we evaluated the ability of the most used severity scoring systems to predict hospital mortality and need for ICU admission in adult, Tunisian injured patients. Four scales were studied and compared: two anatomic scales, the ISS and the NISS, and two physiologic scales, the RTS and the SAPS II. Our results showed that the ISS and the NISS had the best discrimination power associated with a good calibration. Combining the SAPS II scale with the ISS and the NISS did not significantly improve their predictive performance. On comparing the NISS and the ISS, we found that both scores had similar predictive value.

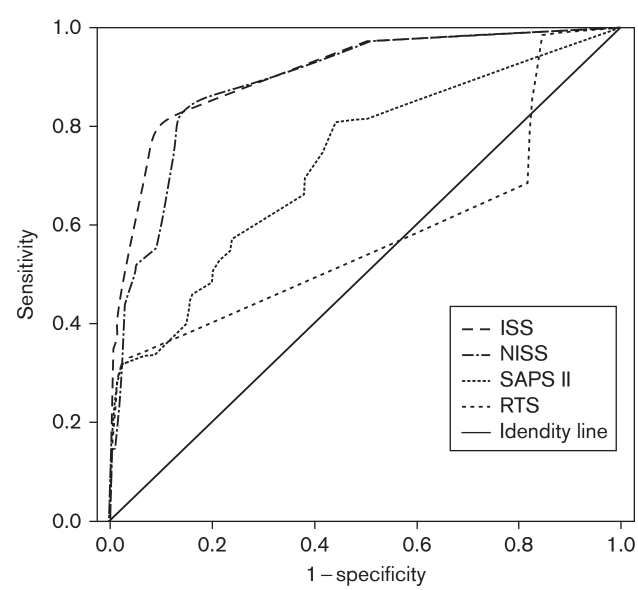
Scoring systems that incorporate physiologic variables are useful in predicting mortality in trauma patients; however, they seem to have lower accuracy than the anatomic-based scoring ones. Bouillon *et al.* [5] compared the ISS and the RTS scores in a population of 2136 German injured patients. They found a better predictive value with the ISS (AUC = 0.96 vs. 0.95). However, some

Fig. 1



Receiver-operating characteristic (ROC) curves for the prediction of intensive care unit admission for Injury Severity Scale (ISS) (area under ROC curve = 0.94), New Injury Severity Scale (NISS) (area under ROC curve = 0.93), Revised Trauma Scale (RTS) (area under ROC curve = 0.70), Simplified Acute Physiology Scale II (SAPS II) (area under ROC curve = 0.73), ISS + SAPS II (area under ROC curve = 0.95), and NISS + SAPS II (area under ROC curve = 0.95).

Fig. 2



Receiver-operating characteristic (ROC) curves for the prediction of hospital mortality for Injury Severity Scale (ISS) (area under ROC curve = 0.91), New Injury Severity Scale (NISS) (area under ROC curve = 0.89), Revised Trauma Scale (RTS) (area under ROC curve = 0.58), Simplified Acute Physiology Score II (SAPS II) (area under ROC curve = 0.73), ISS + SAPS II (area under ROC curve = 0.88), and NISS + SAPS II (area under ROC curve = 0.87).

studies have reported better results for physiologic scales [15–19]. This discordance between results can be mainly attributed to differences in patients' characteristics. In fact, our patients' characteristics were different when compared with those of most published studies. Patients with American Society of Anesthesia >II represented only 17.6% of all our study population. Our relatively lower mortality rate (4.5%) compared with those reported by previous studies [5,7] indicated that our patients probably had better previous health status than in the other studies. In addition, road traffic accidents were less frequent in our study group (27.7%) compared with others [1,2,5,20]. The relatively better outcome observed in our patients compared with those reported by previous studies [5,7] indicated that our patients were less severely injured and, as a consequence, physiologic derangements would be less likely. This would explain partly the superiority of anatomic scales over physiologic ones in our study. Notably, physiologic scales such as SAPS II, APACHE II, and MPM had the best discriminative power when they were applied to the most critically ill patients [21]. As physiologic scores were less sensitive and less specific than anatomic ones, it is not surprising that when combined, the sensitivity and specificity decreased independent of the patient's characteristics. The Trauma and Injury Severity Score that combined ISS and RTS is one example of this nonsynergistic association as we found that RTS showed the lowest performance. The combination of anatomic scales with SAPS II did not improve prediction. In addition, increasing mathematic complexity with mixed scores would limit their clinical use in real practice.

In terms of the choice between the NISS and the ISS, we should note that the results of available studies comparing NISS and ISS with respect to mortality have been contradictory. Several studies have found that equivalence between both scores was mostly observed in patients with low severity of injury [22–24], whereas in patients with higher severity of injury, NISS was found to be superior to ISS [25–27]. Our study findings support this severity-dependent result.

Our study has limitations that should be noted. First, this validation is performed on a single population and the discrimination may not be suitable to other populations. Second, we acknowledge that our patient population was skewed to minor (ISS < 9) injuries. Indeed, many low-severity trauma patients are directly referred from the surrounding areas to our ED, where they received their first assessment. This would explain partly the low rate of patients with severe trauma in the present study. A more reasonable comparison would be to compare those in the more severely injured category (e.g. ISS > 9 or >15) and then compare physiologic and anatomic scales. Unfortunately, the number of such patients was reduced and would not allow such an analysis. Third, functional disability, as a result of trauma, is important in outcome prediction for active individuals. However, we did not assess this criterion in our study. Finally, although many of these scales have limited utility to modify or define patient care in the acute setting such as triage, most of them could be useful to understand variations in outcomes between different centers and countries [28–30].

Conclusion

Our study represents a new contribution toward the validation process of the currently available trauma severity scores outside Europe and North America. We showed the superiority of ISS and NISS over physiologic scales, that is, SAPS II and RTS scores. We also showed that the combination of anatomic with physiologic scales did not improve the prediction performance. This external validation study would provide the possibility of performing outcome comparisons with other centers and develop quality improvement strategies.

Acknowledgements

Conflicts of interest

There are no conflicts of interest.

References

1 Di Bartolomeoa S, Sansonb G, Micheluttoa V, Nardi G, Burba I, Francescutti C, *et al.* Epidemiology of major injury in the population of Friuli Venezia Giulia-Italy. *Injury* 2004; **35**:391–400.
2 Guenther S, Waydhas C, Ose C, Nast-Kolb D. Quality of multiple trauma care in 33 German and Swiss Trauma Centers during a 5-year period: regular versus on-call service. *J Trauma* 2003; **54**:973–978.
3 Draaisma JM, de Haan AF, de Haan AF, Goris RJ. Preventable trauma deaths in The Netherlands. A prospective multicenter study. *J Trauma* 1989; **29**:1552–1557.
4 Van Beeck EF. Injuries in The Netherlands: results of surveillance based public health research. *Eur J Emerg Med* 1999; **6**:389–395.
5 Bouillon B, Lefering R, Vorweg M, Tiling T, Neugebauer E, Troidl H. Trauma score systems: Cologne Validation Study. *J Trauma* 1997; **42**:652–658.
6 Chawda MN, Hildebrand F, Pape HC, Giannoudis PV. Predicting outcome after multiple trauma: which scoring system? *Injury* 2004; **35**:347–358.
7 Baxt WG, Moody P. The differential survival of trauma patients. *J Trauma* 1987; **27**:602–606.
8 Sanddal TL, Esposito TJ, Whitney JR, Hartford D, Taillac PP, Mann NC, Sanddal ND. Analysis of preventable trauma deaths and opportunities for trauma care improvement in Utah. *J Trauma* 2011; **70**:970–977.
9 Pollock DA, O'Neil JM, Parrish RG, Combs DL, Annest JL. Temporal and geographic trends in autopsy frequency of blunt and penetrating trauma deaths in the United States. *JAMA* 1993; **269**:1525–1531.
10 Baker SP, O'Neill B, Haddon W. The Injury Severity Score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma* 1974; **14**:187–196.
11 Osler T, Baker SP, Long WB. A modification to the Injury Severity Score that both improves accuracy and simplifies scoring. *J Trauma* 1997; **43**:922–926.
12 Champion HR, Sacco WJ, Carnazzo AJ, Champion HR, Sacco WJ, Carnazzo AJ, *et al.* Trauma score. *Crit Care Med* 1981; **9**:672–676.
13 Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the trauma score. *J Trauma* 1989; **29**:623–629.
14 Le Gall JR, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American Muticenter study. *JAMA* 1993; **270**:2957–2963.
15 Skaga NO, Eken T, Søvik S, Jones JM, Steen PA. Pre-injury ASA physical status classification is an independent predictor of mortality after trauma. *J Trauma* 2007; **63**:972–978.
16 Gortzis LG, Sakellaropoulos F, Ilias I, Stamoulis K, Dimopoulou I. Predicting ICU survival: a meta-level approach. *BMC Health Serv Res* 2008; **8**:157–164.

17 Yowler C, Cowger J, Barnoski A, Mancuso C, Plaisier B, Super D, Fallon W. Evaluating the process of care for unexpected survivors of trauma using components of TRISS. *Crit Care Med* 1998; **26**:123.

18 Fueglistaler P, Amsler F, Schüepp M, Fueglistaler-Montali I, Attenberger C, Pargger H, *et al*. Prognostic value of Sequential Organ Failure Assessment and Simplified Acute Physiology II score compared with trauma scores in the outcome of multiple-trauma patients. *Am J Surgery* 2010; **200**:204–214.

19 Dossett LA, Redhage LA, Sawyer RG, May AK. Revisiting the validity of APACHE II in the trauma ICU: improved risk stratification in critically injured adults. *Injury* 2009; **40**:993–998.

20 Fischler L, Lelais F, Young J, Buchmann B, Pargger H, Kaufmann M. Assessment of three different mortality prediction models in four well-defined critical care patient groups at two points in time: a prospective cohort study. *Eur J Anaesthesiol* 2007; **24**:676–683.

21 Frankema SP, Steyerberg EW, Edwards MJ, van Vugt AB. Comparison of current injury scales for survival chance estimation: an evaluation comparing the predictive performance of the ISS, NISS, and AP scores in a Dutch local trauma registration. *J Trauma* 2005; **58**:596–604.

22 Nouira S, Belghith M, Elatrous S, Jaafoura M, Ellouzi M, Boujdaria R, *et al*. Predictive value of severity scoring systems: comparison of four models in Tunisian adult intensive care units. *Crit Care Med* 1998; **26**:852–859.

23 Lavoie A, Moore L, LeSage N, Liberman M, Sampalis JS. The Injury Severity Score or the New Injury Severity Score for predicting intensive care unit admission and hospital length of stay? *Injury* 2005; **36**:477–483.

24 Tamin H, Al Hazzouri AZ, Mahfoud Z, Atoui M, El-Chemaly S. The injury severity score or the new injury severity score for predicting mortality, intensive care unit admission and length of hospital stay: experience from a university hospital in a developing country. *Injury* 2008; **39**:115–120.

25 Seow-Yian T, Sloan EP, Zun L, Philip Z. Comparison of the new injury severity score and the injury severity score. *J Trauma* 2004; **55**:1083–1088.

26 Stephenson SCR, Langley JD, Civil ID. Comparing measures of injury severity for use with large databases. *J Trauma* 2002; **53**:326–332.

27 Brenneman FD, Boulanger BR, McLellan A, Redelmeier DA. Measuring injury severity: time for a change? *J Trauma* 1998; **44**:580–582.

28 Lavoie A, Moore L, Lesage N, Liberman M, Sampalis JS. The new injury severity score: a more accurate predictor of in-hospital mortality than the injury severity score. *J Trauma* 2004; **56**:1312–1320.

29 Jamulitrat S, Sangkerd P, Thongpiyapoom S, Na Narong M. A comparison of mortality predictive abilities between NISS and ISS in trauma patients. *J Med Assoc Thai* 2001; **84**:1416–1421.

30 Templeton J, Oakley PA, Mackenzie G, Cook AL, Brand D, Mullins RJ, Trunkey DD. A comparison of patient characteristics and survival in two trauma centres located in different countries. *Injury* 2000; **31**:493–501.

AUTHOR QUERY FORM

LIPPINCOTT WILLIAMS AND WILKINS

JOURNAL NAME: MEJ

ARTICLE NO: 200861

QUERIES AND / OR REMARKS

QUERY NO.	Details Required	Author's Response
	No queries	