A comparison of Injury Severity Score and New Injury Severity Score after penetrating trauma: A prospective analysis

Brian P. Smith, MD, Amy J. Goldberg, MD, John P. Gaughan, PhD,
and **Mark J. Seamon, MD,** *Philadelphia, Pennsylvania*

BACKGROUND: The Injury Severity Score (ISS) has been validated in numerous studies and has become one of the most common trauma scoring systems since its inception. The ISS equation was later modified to create the New Injury Severity Score (NISS). By using the three most severe injuries regardless of body region, the NISS seems well suited to describe patients of penetrating trauma, where injuries often cluster within a single body region. We hypothesized that NISS would better predict outcomes than ISS in penetrating trauma patients.

METHODS: An analysis (June 2008 to March 2009) of all severely injured (length of hospital stay ≥ 48 hours, intensive care unit admission, interhospital transfer, or death) penetrating trauma patients revealed final study sample of 256 patients. ISS and NISS were compared as predictors for both mortality and complications through area under the receiver operating characteristic curve, Hanley-McNeil test, multiple-variable logistic regression, and Hosmer-Lemeshow goodness-of-fit test analysis.

RESULTS: Of 256 study patients, 195 (76.2%) survived until discharge. The mean (ISS, 21.7 ± 21.1 vs. NISS, 27.4 ± 22.0; $p < 0.001$) and median (ISS, 14.0 vs. NISS, 21.0) ISS was lower than those of the NISS. Overall, 173 patients (67.6%) had discordant scores with 26% and 43% having scores greater than 25 (ISS and NISS, respectively, $p < 0.01$). The mortality area under the curve (AUC) for NISS was greater than the AUC for ISS in all penetrating patients (0.930 vs. 0.885, $p = 0.008$), those with penetrating torso injuries (NISS, 0.934 vs. ISS, 0.881, $p < 0.001$), and those with severe (score > 25) injuries (NISS, 0.845 vs. ISS, 0.761, $p < 0.001$). In patients surviving for more than 48 hours, the complications AUC for NISS was also greater than the AUC for ISS (NISS, 0.838 vs. ISS, 0.784; $p = 0.023$).

CONCLUSION: The NISS outperformed ISS as a predictor of both mortality and complications in civilian penetrating trauma patients. These results indicate that NISS is a superior scoring system for patients with penetrating injuries. (*J Trauma Acute Care Surg.* 2015;79: 269–274. Copyright © 2015 Wolters Kluwer Health, Inc. All rights reserved.)

LEVEL OF EVIDENCE: Prognostic study, level III.

KEY WORDS: Trauma scoring system; Injury Severity Score; New Injury Severity Score; trauma performance improvement; trauma benchmarking.

The determination of injury severity is essential to the scientific study of injury, yet the ideal scoring system remains unknown. Trauma scoring systems have several important implications and functions including outcome prediction, triage decision making, risk stratification for research, and benchmarking for performance improvement to allow for better comparison between both individual patients and institutions.[1–20] Multiple scoring systems have been described since 1969[1] when the Association for the Advancement of Automotive Medicine introduced the Abbreviated Injury Scale (AIS), yet each has both strengths and limitations.

Based on the AIS, the Injury Severity Score (ISS) was introduced in 1974 as means of quantifying injury severity.[2] Since then, ISS has been validated in numerous studies and has become one of the most common scoring systems to describe injured patients in trauma outcomes research. However, one well-documented shortcoming of the ISS equation remains—the ISS does not consider multiple severe injuries within the same body region. In response to this limitation, Osler et al.[8] suggested a simple modification of the equation to create the New Injury Severity Score (NISS), using the three most severe injuries regardless of body region.

While all trauma scoring systems have both strengths and limitations, ISS has been proven to underestimate mortality in penetrating trauma patients.[19] The NISS however seems particularly well suited to describe these patients of penetrating trauma, where multiple severe injuries tend to cluster within a single body region. To this end, we hypothesized that NISS would better predict outcomes than ISS would in civilian penetrating trauma patients.

PATIENTS AND METHODS

After institutional review board approval, all penetrating trauma patients at Temple University Hospital who fulfilled at least one of the four criteria were evaluated during the study period (June 2008 to March 2009): length of hospital stay greater than or equal to 48 hours, admission to an intensive care unit, interhospital transfer of a trauma patient, or any death as a

Submitted: March 3, 2015, Revised: April 4, 2015, Accepted: April 7, 2015.
From the Division of Traumatology, Surgical Critical Care and Emergency Surgery (B.P.S., M.J.S.), Hospital of the University of Pennsylvania; and Section of Trauma (A.J.G.), Department of Surgery, Temple University School of Medicine, Philadelphia, Pennsylvania; and Research Institute (J.P.G.), Cooper University Hospital, Camden, New Jersey.
Address for reprints: Mark J. Seamon, MD, Division of Traumatology, Surgical Critical Care and Emergency Surgery, Department of Surgery, Hospital of the University of Pennsylvania, 51 North 39th St, Medical Office Bldg, 1st Floor, Suite 120, Philadelphia, PA 19104; email: mark.seamon@uphs.upenn.edu.

DOI: 10.1097/TA.0000000000000753

result of traumatic injury. Of the 258 evaluated patients, 2 with incomplete autopsy data were excluded, leaving a final study sample of 256 patients for analysis.

Our primary study objective was to compare the ISS and NISS as predictors of hospital mortality. The ISS and NISS were retrospectively calculated by a computer software based on the 2005 version of the Abbreviated Injury Scale (AIS 2005, AAAM Update 2008). AIS data were collected by three trained registrars (Pennsylvania Trauma Systems Foundation) and Collector version 4 Software (Digital Innovation Inc., Forest Hill, MD). The primary study end point, mortality, was defined as death during the index hospital admission. The AIS scores (and subsequent ISS and NISS) for patients who died before surgical intervention were calculated based on autopsy data. The secondary end point, complications, was defined as the occurrence of at least one of the following: acute renal failure, pneumonia, pulmonary embolism, respiratory failure, or sepsis during the index hospitalization in those who survived for more than 48 hours. Secondary hospitalizations were excluded from the analysis. In addition to ISS and NISS, Revised Trauma Scores (RTSs), Trauma and Injury Severity Scores (TRISSs), Penetrating Abdominal Trauma Index (PATI), and Glasgow Coma Scale (GCS) score were calculated for each patient.

The relationship of scoring systems with primary and secondary end points was compared using receiver operator characteristic (ROC) curves. The area under the curve (AUC) was calculated for each scoring system and compared using the Hanley-McNeil test[21] to assess for statistical differences between the AUCs. Because both the ISS and the NISS equations share the highest AIS score in common, the two resulting scores correlate, making them nonindependent. The Hanley-McNeil test[21] is able to differentiate between nonindependent AUC comparisons with increased sensitivity, enabling the detection of statistical differences even if 95% confidence intervals overlap. Logistic regressions were performed using SAS version 9.2 (SAS Institute, Cary, NC) and SPSS version 16.0 (SPSS, Inc., Chicago, IL). All models were tested for calibration using the Hosmer-Lemeshow[22] goodness-of-fit test, a test that assumes the null hypothesis. Thus, for the goodness-of-fit test, a generated $p < 0.05$ indicates poor rather than good model fit, while good model fit is indicated by a $p > 0.05$. Continuous data were reported as medians with value range, and comparisons were made using paired sample *t* tests and Wilcoxon signed-rank tests where appropriate. All tests used two-tailed comparisons, and findings were considered significant (with the exception of the Hosmer-Lemeshow test) if *p* values were <0.05.

RESULTS

Of 256 final study patients, 195 (76.2%) survived to discharge and 61 (23.8%) died. The median age was 26 years (range, 15–78 years; Table 1), and most of the patients were males (91.4%) who sustained gunshot wounds (72.3%). Within the study sample, 94 patients (36.7%) had injuries confined to a single (1 of 6) body region used to calculate the ISS (Fig. 1). Of the 94 patients with single-region injuries, 67 (71.3%) had multiple injuries confined to this single region. One hundred forty-nine patients (58.2%) required surgery for their injuries, and the median hospital length of stay was 5 days (range, 0–69 days).

While the two scoring systems correlated ($r = 0.940$), the mean (ISS, 21.7 ± 21.1 vs. NISS, 27.4 ± 22.0; p <0.001) and median (ISS, 14.0 vs. NISS, 21.0) ISS was lower than those of the NISS. Likewise, there was no difference in the overall range of scores (both 1–75) calculated by ISS and NISS; however, the number of different scores within the range was greater when calculated by NISS (n = 37) than by ISS (n = 31). This is evident in the frequency distribution of the two scoring systems (Fig. 2). Overall, 173 patients (67.6%) had discordant ISS and NISS, and the percentage of patients classified as critical (score > 25) were greater for NISS (43%) than for ISS (26%, $p < 0.01$, Fig. 3).

Comparison of ISS and NISS Predicting Mortality

The primary end point, mortality, occurred in 61 patients (23.8%). The AUC for NISS was greater than the AUC for ISS (0.930 vs. 0.885, $p = 0.008$, Table 2A, Fig. 4) while Hosmer-Lemeshow *p* values for ISS and NISS were 0.057 and 0.190, respectively, indicating goodness of fit. The difference between the areas under ISS and NISS ROC curves was further enhanced by analyzing specific patient subsets including those who sustained penetrating torso injuries (NISS, 0.934 vs. ISS,

TABLE 1. Clinical Characteristics (n = 256)

Patient Characteristics	
Age, y	26, 15–78
Sex, male, n (%)	234 (91.4)
Injury mechanism, n (%)	
Gunshot wound	90 (35.2)
Multiple gunshot wounds	95 (37.1)
Stab wound	48 (18.8)
Multiple stab wounds	23 (9.0)
Initial respiratory rate, breaths/min	19, 0–48
Initial heart rate, beats/min	90, 0–172
Initial systolic blood pressure, mean, mm Hg	111, 0–190
Injury scoring	
ISS	14, 1–75
NISS	21, 1–75
Initial GCS score	15, 3–15
RTS	7.8, 0–7.8
TRISS	0.984, 0–0.995
PATI*	12, 2–37
Outcomes, n (%)	
Operative procedure	149 (58.2)
Complications	25 (9.8)
Pneumonia	3 (1.2)
Sepsis	18 (7.0)
Acute renal failure	3 (1.2)
Pulmonary embolism	2 (0.8)
Ventilator-dependent respiratory failure	5 (2.0)
Ventilator days	0, 0–61
Intensive care unit length of stay, d	0, 0–63
Hospital length of stay, d	5, 0–69
Hospital mortality	61 (23.8)

*Calculated for 76 patients who survived > 48 hours after laparotomy.
Clinical characteristics, injury scoring systems, and outcomes were compared for the study sample.
Continuous variables are reported as median with range.

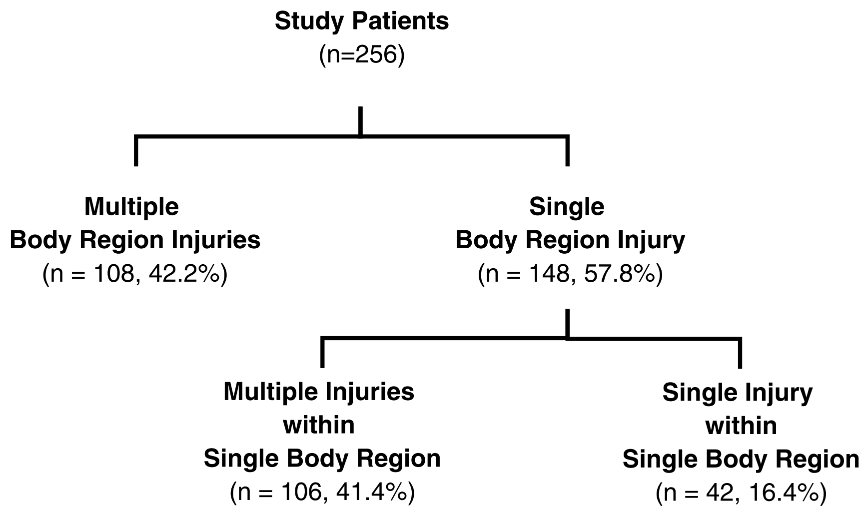

Figure 1. Injuries were characterized with respect to the number of body regions and number of injuries within those regions that were injured.

0.881; $p < 0.001$, Table 2B) and those with severe (score > 25) injuries (NISS, 0.845 vs. ISS, 0.761; $p < 0.001$, Table 2C).

Comparison of ISS, NISS, and PATI Predicting Complications

Thirty-one measured complications occurred in 25 patients (9.7%). Of the 190 patients surviving for more than 48 hours, the area under the NISS ROC curve was greater than that for the ISS curve (NISS, 0.838 vs. ISS, 0.784; $p = 0.023$, Table 2D) when predicting the secondary study end point, complications. In addition, PATI scores were calculated and compared with study scoring systems in the 75 patients who underwent trauma laparotomy for penetrating abdominal injuries. While all tests showed goodness of fit (Table 2E), the

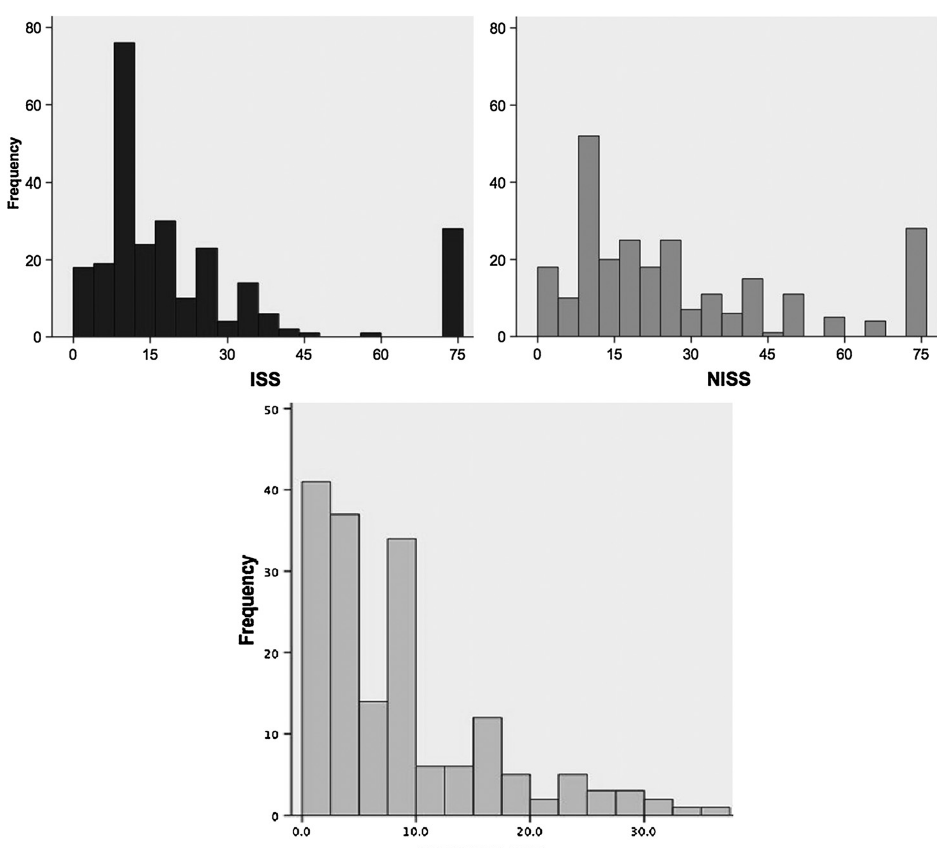

Figure 2. Frequencies for each individual ISS and NISS were assessed as were the difference between the two values (discordance).

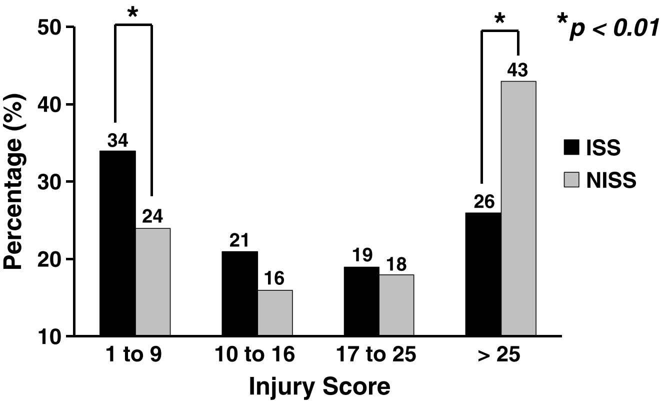

Figure 3. When ISS and NISS were compared based on common categories, NISS less often had minor injury scores (score, 1–9) but more often had critical scores (>25) than ISS.

TABLE 2. Comparison of ISS and NISS Predicting Mortality and Complications

A. ISS/NISS Predict Mortality in Study Sample (n = 256)

Test	AUC	95% CI	HL *p*	Odds Ratio	95% CI	*p*
ISS	0.885	0.833–0.938	0.057	1.100	1.069–1.132	<0.001
NISS*	0.930	0.889–0.971	0.190	1.115	1.085–1.147	<0.001

*Hanley-McNeil test comparing NISS vs. ISS, $p = 0.008$.

B. ISS/NISS Predict Mortality Among Patients With Torso Wounds (n = 171)

Test	AUC	95% CI	HL *p*	Odds Ratio	95% CI	*p*
ISS	0.881	0.820–0.942	0.406	1.093	1.062–1.126	<0.001
NISS*	0.934	0.895–0.973	0.213	1.122	1.083–1.163	<0.001

*Hanley-McNeil test comparing NISS vs. ISS, $p < 0.001$.

C. ISS/NISS Predict Mortality Among Patients With Scores > 25 (n = 89)

Test	AUC	95% CI	HL *p*	Odds Ratio	95% CI	*p*
ISS	0.761	0.670–0.851	0.087	1.062	1.034–1.091	<0.001
NISS*	0.845	0.769–0.921	0.394	1.100	1.061–1.139	<0.001

*Hanley-McNeil test comparing NISS vs. ISS, $p < 0.001$.

D. ISS/NISS Predict Complications After 48 h of Hospitalization (n = 190)

Test	AUC	95% CI	HL *p*	Odds Ratio	95% CI	*p*
ISS	0.784	0.695–0.874	0.329	1.076	1.033–1.120	<0.001
NISS*	0.838	0.765–0.911	0.321	1.080	1.046–1.115	<0.001

*Hanley-McNeil test comparing NISS vs. ISS, $p = 0.023$.

E. ISS/NISS/PATI Predict Complications Among Patients With Abdominal Injuries (n = 75)

Test	AUC	95% CI	HL *p*	Odds Ratio	95% CI	*p*
ISS	0.690	0.557–0.823	0.312	1.071	1.029–1.114	0.001
NISS*	0.805	0.702–0.907	0.258	1.074	1.041–1.108	<0.001
PATI	0.654	0.494–0.815	0.572	1.059	1.012–1.108	0.013

*Hanley-McNeil test comparing NISS vs. ISS, $p = 0.002$.
Hanley-McNeil test comparing NISS vs. PATI, $p = 0.051$.

ISS and NISS were compared through multiple methods, in multiple subsets with both mortality and complications as the studied end points.
CI, confidence interval; HL, Hosmer-Lemeshow.

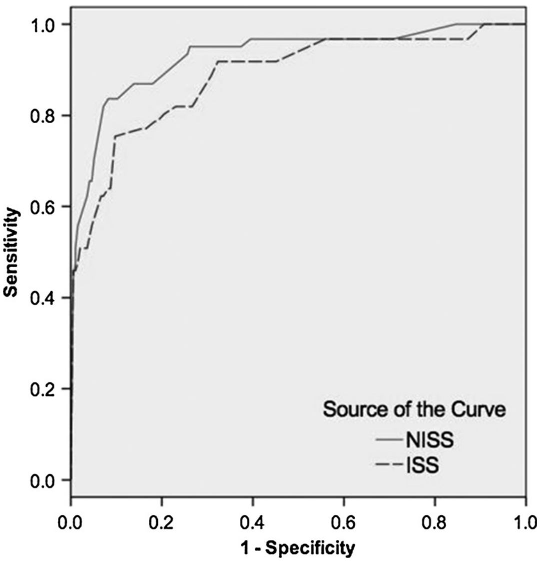

Figure 4. The areas under the ISS and NISS ROC curves were compared for mortality as the end point.

area under the ROC curve for NISS was significantly greater than that for ISS (NISS, 0.805 vs. ISS, 0.690; $p = 0.002$) and greater than that for PATI, although statistical significance was not reached ($p = 0.051$).

DISCUSSION

The NISS outperformed ISS as a predictor of both mortality and complications in this analysis of 256 patients of civilian penetrating trauma. These results indicate that NISS, which considers the three most severe injuries regardless of body region, is a superior scoring system for patients with penetrating injuries.

To our knowledge, this study represents the first comparison of ISS and NISS in a civilian, purely penetrating trauma sample. The NISS seems particularly well suited to model mortality in these patients because penetrating trauma patients commonly sustain numerous severe injuries confined to a single body region. The majority (71%) of our patients with single body region injuries had multiple injuries within that region. While the ISS and NISS scoring systems correlate, they clearly differ for this reason. Our data, like those of other investigators, indicate that discordance between the two systems occurs in 60% of cases.[8,12] Unlike the ISS, the NISS quantifies the three most severe injuries irrespective of region. This allows the NISS to better discriminate each patient from one another, thus increasing the area under the mortality and complication ROC curves compared with those generated by ISS.

Since its inception by Baker et al.[2] in 1974, the ISS has become a validated method of quantifying injury severity and the most common method used to adjust for injury severity both for benchmarking and outcomes research. While its widespread use may be related to its firm entrenchment and the ease with which ISS is calculated, by ignoring concomitant injuries within a single body region, the applicability of ISS to penetrating trauma patients is limited.

The NISS however has been determined to better characterize traumatic injuries and measure injury severity.[8,12,16] Sullivan et al.[11] demonstrated that the two scoring systems were most discordant among pediatric patients with severe (scores > 25) injuries, while Lavoie et al.[12] determined that NISS was a better mortality predictor than ISS for high severity values. In the present study, we have proven that the NISS clearly outperforms the ISS in a cohort of penetrating trauma patients where more than 25% of the sample sustained critical injuries (scores > 25). Our data corroborate the penetrating trauma subset analyses found in two previous reports[11,12] that determined the mortality AUC for the NISS ranged between 0.85 and 0.95. These findings are also concordant with one military report.[10] Although these authors did not demonstrate statistical superiority of the NISS, it should be noted that the methodology by which they compared AUC differed from the present report. In addition, their series included patients of antipersonnel landmines and artillery shells—injury mechanisms that cause a combination of both significant blunt force trauma together with penetrating injury, thereby confounding their scoring system comparisons for penetrating wounds. In a more contemporary comparison, Cook et al.[20] used the National Trauma Data Bank to compare ISS and NISS. While NISS proved superior, the study sample contained only 8% penetrating injury patients. Importantly, the authors found that NISS offers the best rapid injury severity estimate, but the best discrimination and calibration in the large data set was achieved through computer modeling with the Trauma Mortality Prediction Model.

In addition to mortality, we have also determined through multiple variable logistic regression and AUC analysis that NISS was also more effective than ISS in predicting complications 48 hours after penetrating trauma. In those who underwent laparotomy after penetrating abdominal injury, the AUC for NISS was greater ($p = 0.051$) than that of PATI—an anatomic score specifically designed to quantify risk of complication following penetrating abdominal trauma.[3] Similarly, Balogh et al.[13] determined that NISS better predicted postinjury multiple-organ failure than ISS in a mixed trauma population (26% penetrating). Together, these data suggest that the NISS is effective in predicting not only mortality after penetrating injury but also inpatient complications after penetrating injury.

We acknowledge our study limitations. Medical examiner reports were unavailable for 6 of 61 patients who died in this series. While attempts were made to obtain these reports for all patients, operative findings were used for scoring in these six patients. Importantly, our study population was primarily composed of critically injured (24% mortality) gunshot wound patients (72%) who may be more likely to have multiple severe injuries within a single body cavity than stab wound patients with less severe injuries. As a result, our findings should not be generalized to these other penetrating trauma populations. Lastly, we did not evaluate the effect of different AIS triplets on outcomes with identical ISS and NISS—a variable that has been shown to alter mortality predictions.[14,17]

The overlapping confidence intervals of AUC should not be interpreted as a study limitation. Sullivan et al.[11] observed that score correlation, degree of injury severity, and NISS-ISS discordance all contribute to overlap of confidence intervals. In the present study, the correlation of ISS and NISS ($r = 0.94$), high discordance (67.8%), and mortality rates (23.8%) all validate these findings. As the Hanley-McNeil test[21] adjusts for test correlation with the SE of the two measures, statistical significance remains detectable despite the overlapping confidence intervals.

In conclusion, we have determined that NISS outperformed ISS as a predictor for each of the following: mortality after penetrating injury, mortality after penetrating torso injury, mortality after critical (score > 25) penetrating injury, and complications in 48 hours following penetrating injury. These results suggest that NISS, a score that is as simple to calculate as ISS but considers the three most severe injuries regardless of body region, is a superior scoring system for penetrating trauma patients.

AUTHORSHIP

B.P.S. contributed in the study design, data acquisition, data analysis and interpretation, manuscript drafting, and critical manuscript review. A.J.G. contributed in the study design, data analysis and interpretation, and critical manuscript review. J.P.G. contributed in the study design, data analysis and interpretation, and critical manuscript review. M.J.S. contributed in the study design, data acquisition, data analysis and interpretation, manuscript drafting, and critical manuscript review.

DISCLOSURE

The authors declare no conflicts of interest.

REFERENCES

1. States JD. The abbreviated and comprehensive injury scales. In: *STAPP Car Crash Journal*. New York, NY: Society of Automotive Engineers, Inc.; 1969:282–294.
2. Baker SP, O'Neill B, Haddon W Jr, Long WB. The Injury Severity Score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14(3):187–196.
3. Moore EE, Dunn EL, Moore JB, Thompson JS. Penetrating abdominal trauma index. *J Trauma*. 1981;21(6):439–445.
4. Champion HR, Sacco WJ, Hunt TK. Trauma severity scoring to predict mortality. *World J Surg*. 1983;7(1):4–11.
5. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. *Crit Care Med*. 1985;13(10):818–829.
6. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. Trauma Score and the Injury Severity Score. *J Trauma*. 1987;27(4):370–378.
7. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. A revision of the Trauma Score. *J Trauma*. 1989;29(5):623–629.
8. Osler T, Baker SP, Long W. A modification of the Injury Severity Score that both improves accuracy and simplifies scoring. *J Trauma*. 1997;43(6):922–925.
9. Vassar MJ, Lewis FR Jr, Chambers JA, Mullins RJ, O'Brien PE, Weigelt JA, Hoang MT, Holcroft JW. Prediction of outcome in intensive care unit trauma patients: a multicenter study of Acute Physiology and Chronic Health Evaluation (APACHE), Trauma and Injury Severity Score (TRISS), and a 24-hour intensive care unit (ICU) point system. *J Trauma*. 1999;47(2):324–329.
10. Husum H, Strada G. Injury Severity Score versus New Injury Severity Score for penetrating injuries. *Prehosp Disaster Med*. 2002;17(1):27–32.
11. Sullivan T, Haider A, DiRusso SM, Nealon P, Shaukat A, Slim M. Prediction of mortality in pediatric trauma patients: New Injury Severity Score outperforms Injury Severity Score in the severely injured. *J Trauma*. 2003;55(6):1083–1087.

12. 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(6):1312–1320.
13. Balogh Z, Offner PJ, Moore EE, Biffl WL. NISS predicts postinjury multiple organ failure better than the ISS. *J Trauma*. 2000;48(4):624–627.
14. Kilgo PD, Osler TM, Meredith W. The worst injury predicts mortality outcome the best: rethinking the role of multiple injuries in trauma outcome scoring. *J Trauma*. 2003;55(4):599–606.
15. Tay SY, Sloan EP, Zun L, Zaret P. Comparison of the New Injury Severity Score and the Injury Severity Score. *J Trauma*. 2004;56(1):162–164.
16. 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(3):596–604.
17. Aharonson-Daniel L, Giveon A, Stein M, Peleg K; Israel Trauma Group (ITG). Different AIS triplets: different mortality predictions in identical ISS and NISS. *J Trauma*. 2006;61(3):711–717.
18. Nogueira Lde S, Domingues Cde A, Campos Mde A, Sousa RM. Ten years of New Injury Severity Score (NISS): is it a possible change? *Rev Lat Am Enfermagem*. 2008;16(2):314–319.
19. Rowell SE, Rowell SE, Barbosa RR, Diggs BS, Schreiber MA, Holcomb JB, Wade CE, Brasel KJ, Vercruysse G, MacLeod J, et al.; Trauma Outcomes Group. Specific Abbreviated Injury Scale values are responsible for the underestimation of mortality in penetrating trauma patients by the Injury Severity Score. *J Trauma*. 2011;71(2 Suppl 3):S384–S388.
20. Cook A, Weddle J, Baker S, Hosmer D, Glance L, Friedman L, Osler T. A comparison of the Injury Severity Score and the Trauma Mortality Prediction Model. *J Trauma Acute Care Surg*. 2014;76(1):47–52.
21. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983; 148(3):839–843.
22. Lemeshow S, Hosmer DW Jr. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol*. 1982; 115(1):92–106.