Trauma Scoring Systems: A Review

Christopher K Senkowski, MD, Mark G McKenney, MD, FACS

Since West and colleagues[1] showed clear benefits in outcomes for patients treated at specialized trauma centers in the 1970s, patients could no longer be simply transported to the nearest hospital. Scoring systems were initially created for the purposes of field triage.[2] Of necessity these systems must be straightforward and user-friendly for prehospital personnel. Scoring systems should accurately assess severity of injury both anatomically and physiologically. The mechanism of injury is critical. Comorbid factors, age, and clinical judgment also factor into the accuracy of field triage systems. With all these factors incorporated, a scoring system should reliably predict injury severity and patient outcomes.

Beside field triage, scoring systems have found a number of other uses.[3] Because large numbers of patients are quantifiable by scoring systems, these data can be used for quality assurance.[3] Review of records may provide details of proper care, possible areas of preventable morbidity and mortality, and treatment center specific deficiencies or strengths.[4]

Another area where scoring systems have proved valuable is in evaluating trauma care delivery and trauma research.[5] By providing a quantifiable number for groups of trauma patients, comparisons are possible. Researchers can compare different hospitals, different regions, different practice environments, and different modes of therapy. It has become standard in all forms of trauma research to include an injury severity score in the data collection. Scoring systems can also aid in determining entry criteria for prospective research protocols. Using these systems for research has greatly advanced communication among trauma surgeons, health care workers, and researchers by enabling them to speak in similar terms.

Last, trauma scoring systems have the potential to be used in reimbursement assessment. It is generally recognized that trauma and critical care are under-reimbursed.[6] So, although the thought of controlled reimbursement is anathema for most, the era of cost-contained health care delivery is here to stay, and if a quantifiable system proves reliable, it may be that health care regulators should use it.

What follows is a discussion of the current trauma severity scoring systems, and their areas of strength, weakness, and applicability.

GLASGOW COMA SCALE

Developed in 1974 by Teasdale and Jennett[7] from the University of Glasgow, Scotland, the Glasgow Coma Scale (GCS) was the first attempt to quantify severity of head injury. The scale included assessment of three variables (Table 1). The authors chose best motor response to reflect level of CNS function, best verbal response to reflect CNS integrative ability, and eye opening to reflect brainstem function. The admission GCS is predictive of severity of injury.[8] It is used as an initial assessment tool and also in continual reevaluation of head-injured patients. The GCS has become an essential component of other trauma severity systems. The strength of this system lies in that it reliably predicts outcomes for both diffuse and focal lesions.[8] Note that pupillary evaluation is not in-

No competing interests declared.

Some introductory material in this article appears in Shatz D, Kirton OC, McKenney MG, Civetta JM. *Manual of Trauma and Emergency Surgery*, Philadelphia: WB Saunders (October).

Received December 21, 1998; Revised July 21, 1999; Accepted July 22, 1999.
From the Department of Surgery, Mercer University School of Medicine, Savannah, GA (Senkowski) and the Department of Surgery, University of Miami School of Medicine, Miami, FL (McKenney).
Correspondence address: Christopher K Senkowski, MD, PO Box 23089, Savannah, GA 31403.

Published by Elsevier Science Inc.

ISSN 1072-7515/99/$21.00
PII S1072-7515(99)00190-8

Table 1. Glasgow Coma Scale

Best motor response	M score	Best verbal response	V score	Eye opening	E score
Moves limb to command	6	Oriented	5	Spontaneous	4
Localizes to painful stimulus	5	Confused	4	Open to speech	3
Withdraws from painful stimulus	4	Inappropriate words	3	Open to pain	2
Abnormal flexion response	3	Incomprehensible words	2	None	1
Abnormal extension response	2	No verbal	1		
No motor response	1				

GCS = M score + V score + E score. Severe injury, GCS less than 8; moderate, GCS 9 to 12; minor GCS 13 to 15.
From: Teasdale G, Jennett B. Assessment of coma and impaired consciousness. A practical scale. Lancet 1974;2:81, with permission.

cluded in the score because it is not a measure of consciousness.

In a prospective multicenter study of head injured patients with an admission GCS of 9 or less from Los Angeles, Scotland, and The Netherlands,[8] GCS score correlated with mortality despite regional differences in volume, mechanism of injury, and treatment. It is interesting to note that the GCS was not intended initially as a prehospital index. In fact, field GCS scores do not predict outcomes as accurately as admission GCS scores.[9] This effect is largely from initial resuscitation, which can markedly improve initial GCS score.[9] Taken together, the field and admission GCS scores are valuable pieces of information in the treatment and triage of the injured patient.

The GCS does not take into account focal or lateralizing signs, diffuse metabolic processes, or intoxication. This is a potential weakness of the predictive ability of the GCS. Ross and coworkers[10] have recently published data advocating the use of only the best motor response component of the GCS. In their retrospective review of 1,410 patients with prehospital GCS data, the best motor score predicted severe head injury and risk of death from head injury as well as GCS did.

TRAUMA SCORE AND REVISED TRAUMA SCORE

In 1981, Champion and associates[11] published the Trauma Score (TS) as a system for field triage. At the outset, the authors hypothesized that most early trauma deaths were secondary to injury to one or more of three systems: CNS, cardiovascular system, and respiratory system. They next analyzed a large number of variables representing the functional status of these three systems against a cohort of 1,084 patients to select the most independent predictors of outcomes. The resulting TS included five variables: GCS, respiratory rate, respiratory expansion, systolic blood pressure, and capillary refill. The field TS was found to accurately predict survival outcomes in both blunt and penetrating injury.[12] Additionally, it was shown to have strong interrater reliability.[13] In 1989, the same authors reevaluated their system and created the Revised Trauma Score (RTS).[14] In this system, capillary refill and respiratory expansion were dropped because these were often difficult to assess in the field (particularly at night) and had a wide margin for interpretation. There were also concerns that the TS underestimated the severity of head injury in certain instances.[14] The RTS defines three variables: GCS, respiratory rate, and systolic blood pressure. A coded value from 0 to 4 is assigned for each variable (Table 2). From these three coded values a score is generated. It is interesting and important to note that heart rate is not a predicative variable. An RTS score can range from 0 to 12 with lower scores representing increasing severity.

In revising their original scale in 1989, Champion and coworkers,[14] tested the reliability of RTS against two large databases. The Washington Hospi-

Table 2. Revised Trauma Score

Coded value	Glasgow Coma Scale	Systolic blood pressure	Respiratory rate
4	13–15	>89	10–29
3	9–12	76–89	>29
2	6–8	50–75	6–9
1	4–5	1–49	1–5
0	3		

Adapted from: Champion HR, Sacco WJ, Copes WS, et al. A revision of the trauma score. J Trauma 1989;29:624, with permission.

Table 3. Predicting Mortality with Revised Trauma Score

Revised Trauma Score	Mortality (%)
12	<1
10	12
8	33
6	37
4	66
2	70
0	>99

Adapted from: Champion HR, Sacco WJ, Copes WS, et al. A revision of the trauma score. J Trauma 1989;29:625, with permission.

Table 4. Abbreviated Injury Scale

AIS value	Injury severity
0	No injury
1	Minor
2	Moderate
3	Severe (not life-threatening)
4	Severe (life threatening, survival probable)
5	Critical (survival uncertain)

AIS, Abbreviated Injury Scale.
Adapted from: Committee on Medical Aspects of Automotive Safety. Rating the severity of tissue damage. I. The abbreviated scale. JAMA 1971;215:277, with permission.

tal Center database (the database of the principal author) containing 2,166 patients and the Major Trauma Outcome Study (MTOS) database (as developed by the American College of Surgeons Committee on Trauma) containing 26,000 patients were evaluated by RTS. Results showed that an RTS ≤ 11 accurately identified 97.2% of the fatally injured and most of the severely injured as determined by regression analysis. Of 264 false-negative patients (RTS = 12 and severe injury), 167 had severe single body area injuries. This is a potential weakness of RTS used alone and suggests the need for supplementation with anatomic criteria. The decision to transfer a patient to a trauma center based on an RTS ≤ 11 provided a specificity of 82% while maintaining a sensitivity of 59%. Mortality as related to RTS is presented in Table 3.

The RTS is the most widely used prehospital field triage tool, and it has stood the test of time.

INJURY SEVERITY SCALE

The first significant scoring system to be based primarily on anatomic criteria was developed in 1974 by Baker and colleagues[15] from Johns Hopkins University. The Injury Severity Scale (ISS) was created to define injury severity for comparative purposes. It is not a field triage system. The best application of ISS comes in providing researchers a control of the variability of trauma severity for evaluating outcomes. Before this system was developed, it was exceedingly difficult for surgeons to judge the efficacy of treatment of trauma victims.

The strength of this system lies in its incorporation of anatomic indices and severity indices. These authors began by grouping patients according to injury severity using the Abbreviated Injury Scale (AIS).[16] The AIS was developed in 1974 by the American Medical Association committee on Medical Aspects of Automotive Safety. The original aim of this committee, led by Dr Lawrence State, an orthopaedic surgeon, was to provide safety data to automotive design engineers. This committee defined nine categories of severity for several anatomic areas, of which five are applicable to ISS. The 1 to 5 value assignments are presented in Table 4. An AIS of 6 represents a fatal injury for that anatomic area and automatically scores the ISS as 75 (the highest possible score) regardless of any other injuries. AIS scores of 7 to 9 each define fatal injuries at the scene or within 24 hours irrespective of injury severity, so they were not clinically useful in the development of the ISS. These codes were not used in the creation of ISS because meaningful death rates would be obscured by these categories. Next, the authors described the levels of injury by anatomic area (general, head and neck, chest, abdominal, extremity). An example is provided in Table 5.

Baker and associates[15] evaluated 2,128 victims of motor vehicle crashes over a 2-year period in Baltimore. For each patient, the anatomic areas with the highest AIS scores were tabulated. In analyzing these data, a nonlinear relationship was found, such that mortality increased disproportionately with AIS rating of the most severe grade. Also, patients with similar scores sometimes differed greatly in injury severity. For example, a patient with a ruptured spleen (AIS 4) and a pneumothorax (AIS 3) will have different survival probability, compared with a patient with an aortic rupture (AIS 5) and a rib fracture (AIS 2), even though they both have a total score of 7. So, a linear equation would not predict outcomes. Because the simplest nonlinear relationship in mathematics is quadratic, they applied this model to the data and found better correlation of severity and mortality. By taking the sum of the squares of the three highest AIS scores, the best correlation was achieved. Three AIS scores were chosen because us-

Table 5. Abbreviated Injury Scale Classification for Abdominal Injury

AIS value	Description	Examples
0	No injury	None
1	Mild	Muscle ache, seatbelt abrasion
2	Moderate	Major contusion abdominal wall
3	Severe (not life-threatening)	Abdominal organ contusion, bladder rupture ureter avulsion, lumbo-sacral spine fracture without neurologic signs
4	Severe (not life-threatening, survival probable)	Minor laceration abdominal organ ruptured spleen, spine fracture with neurologic signs
5	Critical (survival unclear)	Rupture, avulsion, or laceration to abdominal organs or vessels except spleen, kidney, or ureter

AIS, Abbreviated Injury Scale.
Extracted from: Committee on Medical Aspects of Automotive Safety. Rating the severity of tissue damage. I. The abbreviated scale. JAMA 1971;215:277, with permission.

ing four AIS scores was shown not to improve the correlation. So the ISS is calculated using AIS to determine the three anatomic sites of severest injury, squaring the AIS score of each and taking the sum. An example of a patient with a flail chest, closed femur fracture, and ruptured spleen would earn a score of 41.

$$4^2(\text{chest}) + 3^2(\text{femur}) + 4^2(\text{spleen}) = 41$$

It is important to remember that only one AIS score (the highest) is taken from any single anatomic area, and from these the three highest are used to calculate ISS. This fact impairs the ability of ISS to predict injury severity in patients with multiple injuries in the same anatomic area.

The ISS has become an important predictor of injury severity and mortality.[17] It is the national standard for injury severity assessment. Strengths of ISS are its ability to integrate anatomic areas of injury in formulating a prediction of outcomes. An ISS $\geq$16 has been shown to be associated with a mortality of 10% in a review of 24,192 patients.[18] Ideally, patients with ISS $\geq$ 16 should be treated in centers with experienced trauma personnel. But the ISS is not a prehospital triage tool and in one study calculating ISS in the emergency room, it was found to be unreliable.[19] During initial evaluation and resuscitation, all the exact anatomic injuries are not known. Often the extent of injury and the ISS cannot be truly calculated until operation or extensive diagnostic testing has been performed. The applicability of ISS comes in retrospective analysis of treatment quality and effectiveness, and triage accuracy.[20]

In addition to its weakness as a field triage tool, the initial ISS study evaluated only blunt trauma victims. The 1990 revision of AIS (AIS90) has added significant improvements to the original scale.[21] Specifically, it has (1) expanded the original 75 injuries in 1971 to now more than 2,000 injuries of both blunt and penetrating types; (2) described penetrating injuries similarly across all body areas; (3) revised certain categories to improve prediction in pediatric patients; and (4) greatly expanded descriptions of brain injuries and external injuries.[20] As a result of these improvements in AIS, the ISS has improved its ability to quantify injury severity in both penetrating and bluntly injured patients.[22]

A continuing weakness for the ISS is in predicting outcomes for patients with severe single body area injury. For example in the ISS system, a patient with a gunshot wound injuring the kidney, duodenum, vena cava, and pancreas will get a single AIS score for the abdomen that obviously shortchanges the severity of injury.

In response to this criticism, a new injury severity system (NISS) has been recently proposed by Osler and associates.[23] These authors propose to delineate and code all injuries using AIS90 and then simply take the three highest scores regardless of anatomic area for calculating NISS (sum of the squares). They believe that by taking the three highest scores without regard to anatomic area, the previously described weakness is overcome, and the system provides more accurate predictive ability. Additionally, the calculation is made simpler in NISS. A comparison of ISS with NISS in two separate trauma databases yielded improved prediction of outcomes for NISS in both penetrating and blunt injuries.[23] Brenneman and colleagues[24] further evaluated NISS in a group of 2,328 consecutive blunt trauma admissions in a 4-year period. The authors found average ISS of 25 $\pm$ 13 versus NISS of 33 $\pm$ 18. Scores were discordant in 68%

Table 6. Pediatric Trauma Score

	Category		
Component	+2	+1	−1
Size	≥20 kg	10–20 kg	<10 kg
Airway	Normal	Maintainable	Unmaintainable
Systolic blood pressure	≤90 mmHg	90–50 mmHg	<50 mmHg
CNS	Awake	Obtunded/LOC	Coma or decerebrate
Open wound	None	Minor	Major or penetrating
Skeletal	None	Closed fracture	Open or multiple fracture

LOC, loss of consciousness.
From: Tepas JJ, Mollitt DL, Talbert JL, Bryant M. The pediatric trauma score as a predictor of injury severity in the injured child. J Pediatr Surg 1987;22:15, with permission.

of patients, and NISS provided a more accurate prediction of short-term mortality.

PEDIATRIC TRAUMA SCORE

The number one cause of death in the American pediatric population is trauma.[25] Most of the field triage tools are not applicable for pediatric trauma victims. For example, normal respiratory rate, heart rate, and systolic blood pressure vary considerably with infancy and childhood. Additionally, the verbal response as used in GCS is obviously inaccurate for young children. For these reasons, Tepas and colleagues[26] created the Pediatric Trauma Score (PTS). In their scale, six variables are included (Table 6). Each variable is scored +2 for minimal or no injury, +1 for minor or potentially major injury, or −1 for major or life-threatening injury. The total score ranges from +12 to −6 with increasing severity.

Specifically, the authors reasoned that smaller sized children had less physiologic reserve, so weight became a variable. Systolic blood pressure, airway status, and level of consciousness were variables included similar to adult scoring systems. Presence of open wounds or fractures were the final two variables. The presence of these injuries suggests severe energy transfer and positively correlates with concomitant visceral injury.[25]

These authors then correlated PTS with ISS at discharge or autopsy in 230 patients from two separate databases.[26] The PTS was a reliable predictor of severity and outcomes as judged by ISS. A followup study recommended a PTS ≤ 8 (ie, 8 to −6) as identifying a cohort of patients with an increased risk of mortality.[27]

Critics of the PTS claim that the additional training of prehospital personnel to use this system is unnecessary because existing adult systems (specifically RTS) work quite well when applied to the pediatric population.[28] In fact, two reviews have shown RTS to be as effective as PTS in guiding prehospital triage of the injured pediatric patient.[29,30] Specifically, in a 1990 study published in JAMA by Kaufmann and coworkers, the PTS was of no statistical advantage as compared with the RTS. This study reviewed 376 patients and found triage accuracy rates of 68.3% for PTS and 78.8% for RTS.[31]

OTHER SCORING SYSTEMS

A number of other scoring systems deserve mention both for completion and because, although not universally adopted, each has contributed to the overall understanding of trauma triage and injury severity. The Triage Index[32] and the Illness-Injury Severity Index[33] were both excellent early attempts to combine physiologic and anatomic indices for quantifying injury severity that unfortunately never caught on. The CRAMS (circulation, respiration, abdomen, motor, speech)[34] and Trauma Triage Rule[35] were both elegant in their simplicity but could not be consistently verified on repeat scrutiny.[36] Lastly deserving of mention is the Revised Estimated Survival Probability Index[37] that attempted to use all patient injuries using ICD coding data to predict outcomes. It was never able to outperform ISS and was abandoned.

TRISS

By combining the anatomic criteria of the ISS with the physiologic criteria of the RTS, the "TRISS method" for analyzing trauma data was elucidated.[38] Using logistic regression analysis, the TRISS method correlates RTS with ISS to create an S_{50} isobar on which a 50% survival is predicted. Patient probability of survival is then plotted on the RTS versus ISS graph (Fig. 1). With the isobar in place, survivors who fall above the isobar (unexpected survivors) and those deaths below the isobar (unexpected deaths)

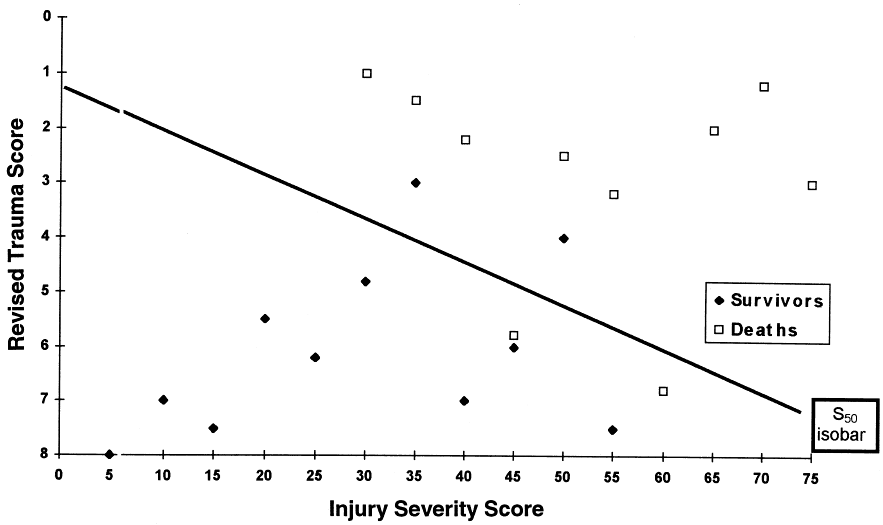

Figure 1. Revised Trauma Score plotted against Injury Severity Score to predict mortality: the "TRISS method."

can be identified. It must be kept in mind that the 50% survivor cutoff is arbitrary and only provides a method for isolating outlying patients in a particular series. They must then be analyzed on an individual basis. For example, a patient with 55% chance of survival as predicted by TRISS who dies will appear as a preventable death when, in fact, that may not be correct.

When TRISS is used to compare trauma patient outcomes data against the baseline database (MTOS), it has been described as PRE evaluation or preliminary outcomes-based evaluation. By this method, any trauma service regardless of size can compare itself to the MTOS database for outcomes analysis. A second and equally important use of TRISS is DEF evaluation or definitive outcomes-based evaluation. In this mode, two patient groups are compared with each other and not to any baseline normative database. Applications of the DEF method include comparisons of treatment protocols or comparisons between treatment centers.

Statistical details of the TRISS method can be daunting. The basic idea is to determine a probability of survival (Ps) for each patient. The formula is as follows:

$$\text{Ps} = 1/(1 + e^{-\beta}).$$

In this formula:

$$\beta = \beta_0 + \beta_1(\text{RTS}) + \beta_2(\text{ISS}) + \beta_3(\text{Age})$$

the β values are derived using logistic analysis of the entire database. In the original TRISS article the β values were calculated from the Major Trauma Outcome Database (Table 7). The RTS score is calculated using weighted coefficients also derived from the MTOS:

$$\text{RTS} = 0.9364(\text{GCS}) + 0.7326(\text{SBP}) + 0.2908(\text{RR}).$$

Going back to the β value formula, the ISS score is inserted as calculated and for Age, a 1 is inserted for patients 55 years or older and a 0 for patients under 55 years. Because survival probabilities are compared with actual outcomes, individual outliers can be identified (both unexpected survivors and unexpected deaths).

When using TRISS for PRE type analysis, the formula coefficients will come from the normative database (usually MTOS) as opposed to DEF type analysis where coefficients are derived from the databases under scrutiny.

The TRISS methodology has been widely used for institutional internal quality assurance (PRE type evaluation).[39–41] With the TRISS graph, preventable deaths can be identified, reviewed, and discussed.

Table 7. TRISS Coefficients

	β_0	β_1(RTS)	β_2(ISS)	β_3(Age)
Blunt	−1.247	0.9544	−0.0768	−1.9052
Penetrating	−0.6029	1.143	−0.1516	−2.6676

ISS, Injury Severity Score; RTS, Revised Trauma Score.
From: Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. J Trauma 1987;27:373, with permission.

Revised Trauma Score	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75
1	0.33	0.25	0.19	0.14	0.1	0.07	0.05	0.03	0.02	0.02	0.01	0.01	0.01	0	0
2	0.56	0.47	0.37	0.29	0.22	0.16	0.11	0.08	0.06	0.04	0.03	0.02	0.01	0	0.01
3	0.77	0.69	0.61	0.51	0.42	0.33	0.25	0.19	0.13	0.1	0.07	0.05	0.03	0.02	0.02
4	0.9	0.86	0.8	0.73	0.65	0.56	0.46	0.37	0.29	0.22	0.16	0.11	0.08	0.06	0.04
5	0.96	0.94	0.91	0.88	0.83	0.77	0.69	0.61	0.51	0.42	0.33	0.25	0.18	0.13	0.09
6	0.98	0.98	0.96	0.95	0.93	0.9	0.85	0.8	0.73	0.65	0.56	0.46	0.37	0.28	0.21
7	0.99	0.99	0.99	0.98	0.97	0.96	0.94	0.91	0.88	0.83	0.77	0.69	0.6	0.51	0.41
8	0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.97	0.95	0.93	0.89	0.85	0.8	0.73	0.65
	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75

Injury Severity Score

Figure 2. TRISS scan. Probability of survival for blunt trauma patients age 54 or less. (Adapted from: Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: the TRISS method. J Trauma 1987;27:377, with permission.)

Additionally, efficacy of new protocols can be readily evaluated with the TRISS methodology (DEF type evaluation). TRISS has also been used to evaluate outcomes among different treatment centers to identify strengths and weaknesses.[42]

To facilitate PRE type comparisons with MTOS, the MTOS database was evaluated by the TRISS method and a TRISSSCAN grid (example, Fig. 2) was created for two age groups (age>55 and age <55).[39] The grid provides a quick survival probability based on ISS and RTS. This grid has been used by some emergency departments to initiate triage of trauma resources.[36]

Weaknesses of TRISS include underestimation of patients with multiple severe injuries to the same anatomic area (inherent in ISS calculation as discussed previously). Also, there is a need to further break down survival probability by age.

ASCOT

In an attempt to further improve TRISS, Champion and associates[43] in 1990, created ASCOT (A Severity Characteristic of Trauma) using the Anatomic Profile.[44] Like ISS, the Anatomic Profile is based on AIS scores with some important differences. The Anatomic Profile created four components: component A included head, brain, and spinal cord injuries; component B included thoracic and anterior neck injuries; component C included all other major injuries; and component D included all minor injuries. All AIS scores > 3 for each component were included to achieve the final score. These AIS > 3 scores were squared and summed. The ASCOT authors found component D was not useful in predicting mortality, so was dropped from the scoring system. In the ASCOT system, age was stratified into five different ranges to provide more useful information (Table 8). By merging age, RTS, and Anatomic Profile into similar logistic regression analysis as used to derive TRISS, the ASCOT instrument emerged. ASCOT calculates probability of survival with the following formula:

$$Ps = 1/(1 + e^{-K}),$$

where

$$K = K_0 + K_1(GCS) + K_2(SBP) + K_3(RR) + K_4(A) + K_5(B) + K_6(C) + K_7(Age)$$

Table 8. ASCOT Patient Age Characterization

Age value	Ages (y)
0	0–54
1	55–64
2	65–74
3	75–84
4	>84

ASCOT, A Severity Characteristic of Trauma.
From: Champion HR, Copes WS, Sacco WJ, et al. A new characterization of injury severity. J Trauma 1990;30:540, with permission.

Table 9. ASCOT Set Aside Groups with Survival Probabilities

	Survivors (%)	
Group	Blunt injury	Penetrating injury
AIS score 6, RTS = 0	0.0	0.0
Maximum AIS score < 6, RTS = 0	1.4	2.6
AIS score 6, RTS > 0	22.9	22.2
Maximum AIS score = 1 or 2, RTS > 0	99.8	99.9

AIS, Abbreviated Injury Scale; ASCOT, A Severity Characteristic of Trauma; RTS, Revised Trauma Scale.
From: Champion HR, Copes WS, Sacco WJ, et al. A new characterization of injury severity. J Trauma 1990;30:541, with permission.

where SBP-systolic blood pressure and RR-respiratory rate. Age is stratified into five groups and each is given a value from 0 to 4 (Table 8). The ASCOT system also selects out those patients with extremely good (highest AIS = 1 or 2) or extremely poor (AIS = 6 or RTS = 0) prognoses. These subgroups are assigned survival probabilities and are exempted from the logistic regression analysis (Table 9).

Proponents of this tool claim it better controls for age and is also more applicable to patients with multiple single body area injuries. There have been three studies comparing ASCOT to TRISS using non-MTOS data.[45-47] Markle and associates[45] studied 5,685 patients from eight hospitals affiliated with the New York Medical College (Institute for Trauma and Emergency Care [ITEC] database) and found that neither TRISS nor ASCOT reliably predicted survival. In this study coefficients were used from the MTOS data. They did find that ASCOT predicted fewer unexpected deaths in patients with CNS injury and also in those patients with multiple injuries to a single anatomic area. Hannan and associates,[46] using the same ITEC database, repeated the analysis with new coefficients derived directly from the ITEC database. They found that ASCOT did predict survival acceptably for blunt trauma patients, but that TRISS did not perform acceptably. The third study using four MTOS data sites found ASCOT outperformed TRISS in both penetrating and blunt trauma patients.[47] But in this third study, the only model that met the statistical requirement of predictive ability was ASCOT for predicting survival in blunt trauma patients (HL test < 15.5, see below). Interestingly, TRISS and ASCOT were equally able to predict outcomes in pediatric patients.[47]

ADDITIONAL STATISTICAL TERMS FOR UNDERSTANDING TRISS AND ASCOT

In addition to grasping the complexities of the logistic regression analysis formulas previously shown, one must also understand a few more terms:

The Z statistic

The Z statistic[48] is used as a general measure of how well two different subpopulations compare. For example, it quantifies the difference between actual and expected survivors or deaths. A positive Z score indicates that there were more actual survivors (or deaths) than expected, and a negative Z score indicates less actual survivors than expected. When studying survivors, a positive Z score is desirable (ie, there were more actual survivors than predicted). Conversely, when studying mortalities, a negative Z score is desirable because this implies more expected death than actual deaths. The formula for computing Z statistic is as follows:

In considering deaths,

$$Z = (D - \Sigma Q_i)/\sqrt{\Sigma P_i Q_i},$$

where D = Actual number of deaths, P_i = predicted survival probability for patient i, ΣQ_i = predicted number of deaths, and Qi = (1 − Pi) predicted probability of death for patient i.

In considering survivors,

$$Z = S - \Sigma P_i/\sqrt{\Sigma P_i Q_i}.$$

If the absolute Z score is greater than 1.96, this indicates that the difference between the test population and the baseline population with respect to expected survivors (or deaths) is significant at the < 0.05 level of significance.

The Receiver-Operating Characteristic curve

The receiver-operating characteristic (ROC)[49,50] curve analysis is a means to measure the power of a test to separate two mutually exclusive subpopulations. The ROC statistic is defined as the area under the graph of (sensitivity) (1-specificity). By plotting the true-positive fraction (sensitivities) against the false-positive fraction (1-specificity), the curve is constructed. It can be understood as an infinite series of likelihood ratios.[50] By assessing the area under the entire curve rather than looking at particular points on the curve, a more accurate assessment of the test can be made. An ROC of 1.0 (the best) means the test will perfectly discriminate between two subpopulations; an ROC of 0.50 (the worst) would indicate that the test performs no better than chance. In a study by Rutledge and colleagues[51] comparing the ability of ISS and TRISS to predict survival, the authors found an ROC of 0.667 for ISS and an ROC of 0.877 for TRISS. Using this example, it would seem that TRISS has an increased ability over ISS to predict survival. For examples of ROC curves, see Figure 3.

The Hosmer-Lemeshow test

This is a test of "goodness of fit." It will compare two populations over a range of 10 intervals (subdivisions) to measure predictive ability. In other words, the Hosmer-Lemeshow (HL) test will compare predicted survival with actual survival over a range of differing injury severities. The HL test serves as a calibrator. At the <0.05 level of significance, an HL test of <15.5 is evidence of adequate fit.

When using this test, the created intervals become critical. Critics have cited this as a weakness of the test.[45,47] Specifically, intervals can be set up in at least three ways: (1) intervals are of specific equal length regardless of case distribution; (2) cases are distributed so that one tenth is in each interval; and (3) intervals are created so that each subdivision has equal number of survivors. Because of this variability, the HL test has been criticized because the investigator can manipulate the "calibrator" to show the data in the best light, and it is unclear in many instances in what way the intervals should be reported.

ICD-9 AND ICISS SYSTEMS

The latest injury severity scoring systems to arrive are the ICD-9–based models. The idea to use ICD-9 coded data to create an injury severity scoring system was a response to the amount of resources needed to implement ISS-based systems. To classify patients by AIS, a dedicated trauma registry and staff to encode, enter, and police the data is required. The AIS90 scoring is complicated and requires significant knowledge and training. Many major urban trauma centers have dedicated staff who function to create and maintain trauma registries. Smaller urban and rural centers rarely have these resources. So these hospitals are unable to use TRISS and ASCOT for quality assurance, and little data are available from these centers for national examination. Most of our outcomes data results come from large dedicated trauma centers. It would be quite interesting and important to examine similar data from small hospitals that care for the injured patient. Because ICD-9 coding already exists in every hospital regardless of size, a model based on it would provide broad availability for outcomes analysis.

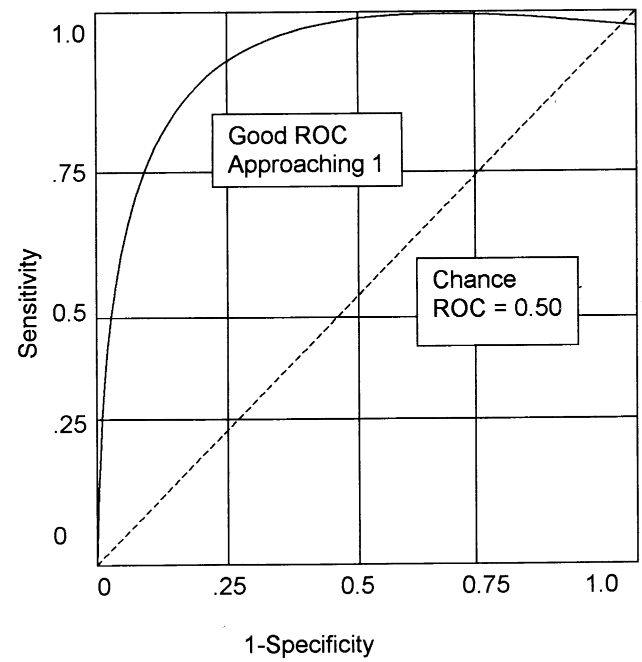

Figure 3. Examples of Receiver-Operating curves. ROC, Receiver-Operating characteristic.

The ICD-9 system is a nomenclature system and not a scoring system, but it does provide anatomic descriptions of injuries, surgical procedures performed, and cause of injury (E code). In 1993, Rutledge and coworkers[53] used the North Carolina Trauma Registry (NCTR) of 37,100 patients to calculate mortality risk ratios (MRRs) for each ICD-9 code. This assigned a relative severity rate to each code. They next compared the ability to predict outcomes of the ICD-9 system with ISS. They concluded (1) ISS was a significant predictor of survival but poorly sensitive when predicting death; (2) the MRR for the primary injury code and the second- and third-coded injuries, the first- and second-coded procedures, and the cause of injury (E code) each independently correlated with survival; and (3) that a multivariate model based on all MRRs (ie, injury codes, procedures codes, and E code) for each patient was a better predictor of outcomes than ISS.[53] This study, along with further work using computer-generated models,[54] provided strong data for the possibility of predicting injury severity and outcomes based on the widely available ICD-9 data.

A major weakness of these studies is that the ISS scores used to compare against ICD-9 predictions were not derived from AIS scores but rather from the ICD-9 codes themselves by a method described by MacKenzie and colleagues.[55] This method inherently lowers the predictive ability of ISS. Critics claim that the mortality rate (2.4%) and the injury severity (1.6 injuries per patient) in the NCTR are too low and are not comparable with other large trauma databases.[53,54] This would allow any system to predict survival with good accuracy (ie, if the model predicts 100% survival in a database with a 2.4% mortality, the system will have a 97.6% accuracy rate). Lastly, the MRR values are derived as if they are independent variables when, in fact, combinations of injuries have different survival probabilities than the probabilities of the individual injuries.

In 1996, Osler and colleagues[56] created ICISS (*International Classification of Diseases*, 9th Edition, Injury Severity Score) as the most recent ICD-9–based scoring system. The authors calculated survival risk ratios (SRRs) using the NCTR for each ICD-9 code. By simply dividing the number of times the code occurred in a surviving patient by the total number of times the code occurred in the database, the SRR was defined. Simply multiplying all the SRRs for a given patient results in the ICISS score.

$$\text{ICISS} = P_{\text{survival injury 1}} + P_{\text{surv inj 2}} + \ldots P_{\text{surv inj last}}.$$

Next, using a New Mexico trauma registry of 3,142 patients, AIS-based ISS scores were compared with ICISS-derived scores.[56] The ICISS was a better predictor of outcomes than ISS (ROC for ISS = 0.872 versus ICISS = 0.921) with better "goodness of fit" (HL test ICISS = 10.6, ISS = 51.3).

The ICISS system has yet to be rigorously tested but holds promise as a tool for evaluating injury severity. Weaknesses are in the assumption that each ICD-9 code is independent in calculating SRR (ie, that some injuries in combination are more severe than the simple addition of the individual injuries). Additionally, trivial injury severity will be universally overestimated because any mortality would not likely be from the trivial injury but rather a concomitant more severe injury. Finally, the SRRs were calculated in some instances with very small (as little as two) cohorts of patients. The authors themselves cite these weaknesses and caution against declaring one system superior to another and that further research is critical and necessary.[56]

OVERTRIAGE AND UNDERTRIAGE

All scoring and triage systems are susceptible to the problem of overtriage and undertriage. Overtriage is defined as the number of patients with minor injuries who are transported to a specialized trauma center. Undertriage is defined as the number of patients with severe injuries who were inappropriately triaged to nontrauma centers. There is general agreement that efforts to minimize undertriage to less than 5% to 10% at the expense of increased overtriage is desirable.[57,58] The American College of Surgeons (ACS) Committee on Trauma estimated an overtriage rate of 50% would provide acceptable sensitivity and minimal undertriage.[59]

The adverse effects of undertriage are obvious; patients with severe injuries not arriving at designated trauma centers may have increased morbidity and mortality. The adverse effects of overtriage include overburdening a specialized center with patients not needing the resources. It is expensive and fatiguing for the hospital and staff.

To minimize overtriage without compromising sensitivity, a number of trauma surgeons have investigated secondary in-hospital triage mechanisms.[60–63] These usually involve different levels of trauma team activation as coordinated between emergency medicine physicians and surgeons.

In 1993, Phillips and Buchman,[60] from Johns Hopkins University, put forth the idea of a "two-tiered" trauma team response algorithm. The ACS triage criteria as presented in Fig. 4 were further subdivided into Delta and Echo classifications. The Delta criteria included all physiologic indicators of severe injury and many mechanisms of injury indicators (ie, all shotgun injuries). Any Delta criteria would activate the full trauma team response. Patients not fulfilling any Delta criteria were placed in the Echo category and received "expedited trauma care." Using this system, under- and overtriage rates were 5% and 32%, respectively.[60]

DeKeyser and colleagues,[61] from Virginia, stratified their trauma admissions into code blue and code yellow categories. Code blue was given to patients with significant physiologic impairment (airway compromise, hypotension, GCS less than 8) or if penetrating injury to head, neck, torso, or crush injury were present. All others received a code yellow. Code blue activated a full 16-member response, whereas code yellow summoned 8 members. The division was based on the fact that the physiologic cri-

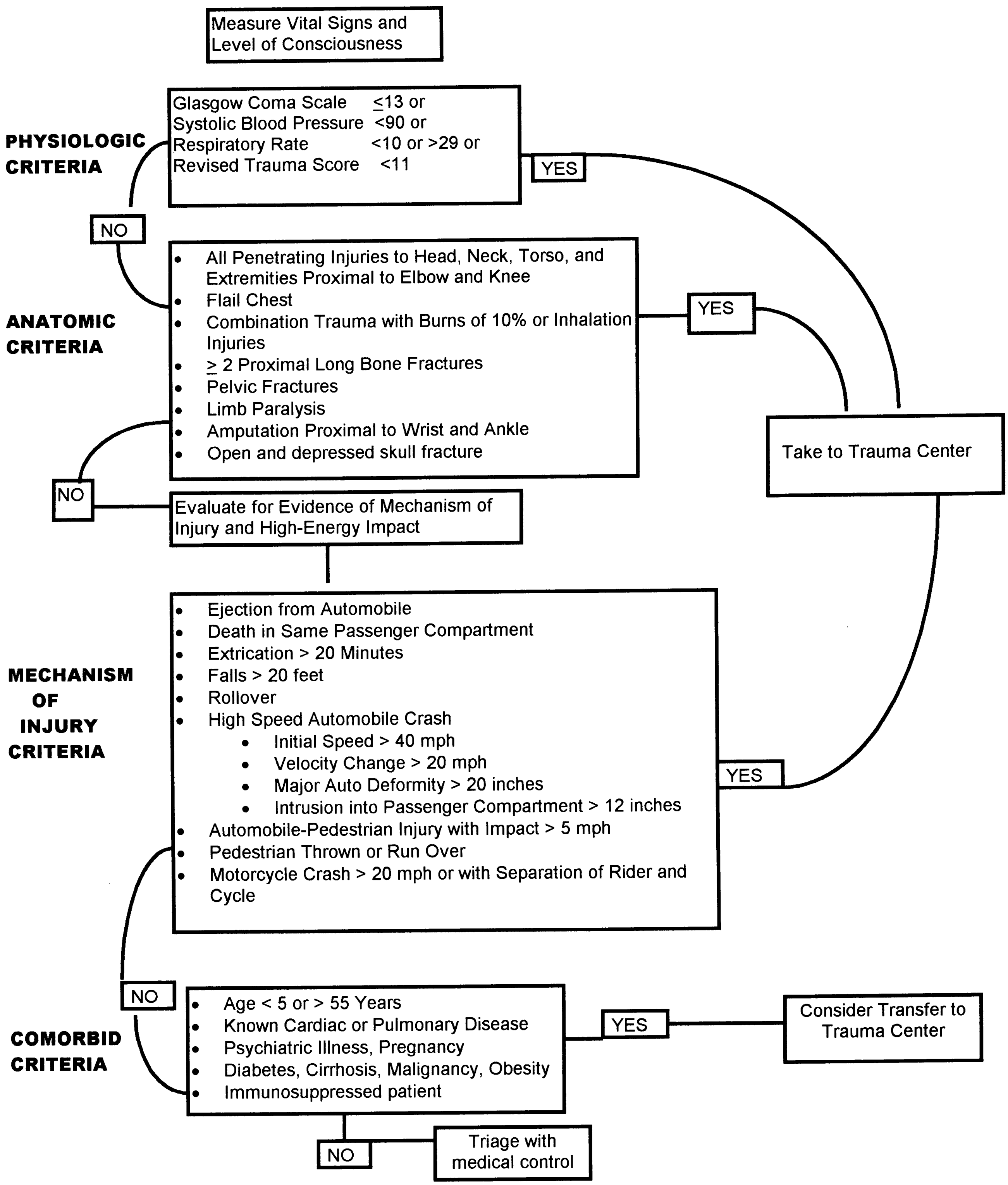

Figure 4. Triage algorithm. (From: American College of Surgeons Committee on Trauma. Resources for optimal care of the injured patient. Chicago: American College of Surgeons; 1999, with permission.)

teria are highly specific (ie, when positive, these patients are indeed severely injured) (98%).[14] Because physiologic criteria are not very sensitive (ie, if negative, the patient may still be severely injured), the added evaluation with anatomic and mechanism of injury criteria would be applied. Cottington and coworkers[62] concurred in their study that along with physiologic criteria, penetrating trauma, chest injury, and abdominal injury, each increased overall specificity. The code blue and yellow system yielded significant cost savings, but the authors did not specifically address patient outcomes in terms of preventable deaths or delays in treatment with this system.

A review of a Washington DC Level I trauma center by Ochsner and colleagues[63] revealed 30% to 32% of patients were discharged in 24 hours. Additional scrutiny showed that these patients were most often overtriaged for mechanism of injury criteria. A phase 1 study used ACS criteria for full trauma alert. After evaluation of the results, two criteria accounted for 64% of overtriage: (1) penetrating gunshot, stab, or ice pick wound to head, neck, trunk, or groin or to an extremity with obvious bleeding or fracture; and (2) head trauma with $GCS < 13$ but > 11 (ie, $GCS = 12$). In phases II and III of the study, the authors created two levels of trauma alerts.[63] The full trauma alert criteria were narrowed by triaging stab wounds and open fractures to a second tier and decreasing GCS criteria for full activation to ≤ 11. The final criteria provided appropriate triage for 383 (79%) of 486 patients. There was an overtriage rate of 12% and an undertriage rate of 9%.[45] The authors were also able to demonstrate that "two tiers" of in-hospital trauma response were safe and cost-effective. In summary, the ACS criteria (Fig. 4) can provide initial direction of flow of trauma patients from the field with acceptable undertriage rates, and in-hospital systems can be individually constructed to lessen overtriage rates.

POPULATION DATA SETS

Predictions of injury severity are only as good as the data entered and are not necessarily always representative of the patient. It is paramount to understand that prediction of outcomes cannot apply to clinical decision-making for an individual patient. These survival probabilities and triage criteria are based on the evaluation of large data sets. It is inappropriate to predict outcomes of the individual patient or to initiate treatment based on predicted outcomes. A clinician must understand that a system with a 5% misclassification error applies to the whole set and that select groups of complicated patients may have much higher misclassification rates (up to 40%).[56] Although these scoring systems enable caregivers to make clinical decisions in regard to a population, they cannot be used on the individual patient in isolation.

In conclusion, the field of trauma scoring has exploded in the last 10 years, with exciting results. Trauma systems are well implemented in many areas of the country, rigorously tested, and continually improved and revised. As trauma caregivers, we are developing better prehospital triage guidelines based on refined and accurate trauma registries. We have discussed the GCS and the RTS and recommend these as the most reliable prehospital triage instruments. We have created systems for assessing outcomes to continue the process of improving rendered care. These quality improvement instruments include ISS, NISS, TRISS, ASCOT, and ICISS, and all are vying for the title of the best tool. Currently, ISS is embedded into many trauma registries, and as a result is linked to many years of data. It is unlikely that there is one "best" tool but the debate continues. It is important to understand each of these systems to discuss research ideas and beliefs in the same language.

References

1. West JG, Trunkey DD, Lim RC. Systems of trauma care: A study of two counties. Arch Surg 1979;114:445–460.
2. Maslanka AM. Scoring systems and triage from the field. Emerg Med Clin North Am 1993;11:15–27.
3. Wisner DH. History and current status of trauma scoring systems. Arch Surg 1992;127:111–117.
4. MacKenzie EJ. Injury severity scales: Overview and directions for future research. Am J Emerg Med 1984;2:537–549.
5. Mendeloff JM, Cayten CG. Trauma systems and public policy. Annu Rev Publ Health 1991;12:401–424.
6. Schwab CW, Young G, Civil I, et al. DRG reimbursement for trauma: The demise of the trauma center. J Trauma 1988;28:939–946.
7. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: A practical scale. Lancet 1974;2:81–84.
8. Langfitt TW. Measuring the outcome from head injuries. J Neurosurg 1978;48:673–678.
9. Jennet B, Teasdale G, Braakman R, et al. Prognosis of patients with severe head injury. Neurosurgery 1979;4:283–289.
10. Ross SE, Leipold C, Terregino C, et al. Efficacy of the motor component of the Glasgow Coma Scale in trauma triage. J Trauma 1998;45:42–44.
11. Champion HR, Sacco WJ, Carnazzo AJ, et al. The trauma score. Crit Care Med 1981;9:672–676.
12. Sacco WJ, Champion HR, Gainer P, et al. The trauma score as applied to penetrating trauma. Ann Emerg Med 1984;13:415–418.
13. Moreau M, Gainer PS, Champion HR, et al. Application of the trauma score in the prehospital setting. Ann Emerg Med 1985;14:1049–54.
14. Champion HR, Sacco WJ, Copes WS, et al. A revision of the trauma score. J Trauma 1989;29:623–629.

15. Baker SP, O'Neill B, Haddon W, et al. The injury severity score: A method for describing patients with multiple injuries and evaluating emergency care. J Trauma 1974;14:187–196.
16. Committee on Medical Aspects of Automotive Safety. Rating the severity of tissue damage: The Abbreviated Scale. JAMA 1971; 215:277–280.
17. Trunkey D. Panel: Current status of trauma severity indices. J Trauma 1983;23:185–188.
18. Champion HR, Frey CF. Report on the major trauma outcome study. Presented at the American College of Surgeons Committee on Trauma. Chicago;1986.
19. MacKenzie EJ, Shapiro S, Eastham J. Rating AIS severity using emergency department sheets versus inpatient charts. J Trauma 1985;25:984–988.
20. Trunkey DD. Overview of trauma. Surg Clin North Am 1982;62:3–7.
21. American Association for Automotive Medicine. The abbreviated injury scale (AIS)—1990 revision. Des Plaines, IL;1990.
22. Champion HR, Copes WS, Sacco W, et al. Research in trauma outcomes using the MTOS database. AHCPR Grant HS 06721, 1994.
23. Osler T, Baker SP, Long W. NISS: The new injury severity score. Presented at the Eastern Association for the Surgery of Trauma. Sanibel, FL;1997.
24. Brenneman FD, Boulanger BR, McLellan BA, et al. Measuring injury severity: time for a change? J Trauma 1998;44:580–582.
25. Haller JR. Pediatric trauma: The number 1 killer of children. JAMA 1983;249:47–49.
26. Tepas JJ, Mollitt DL, Talbert JL, et al. The pediatric trauma score as a predictor of injury severity in the injured child. J Pediatr Surg 1987;22:14–18.
27. Tepas JJ, Ramenofsky ML, Mollitt DL, et al. The pediatric trauma score as a predictor of injury severity: An objective assessment. J Trauma 1988;28:425–427.
28. Nayduch DA, Moylan J, Rutledge R, et al. Comparison of the ability of adult and pediatric trauma scores to predict pediatric outcome following major trauma. J Trauma 1991;31:452–458.
29. Eichelberger MR, Goschali CS, Sacco WJ, et al. A comparison of the Trauma Score, the Revised Trauma Score, and the Pediatric Trauma Score. Ann Emerg Med 1989;18:1053–1058.
30. Kirkpatrick JR, Youmans RL. Trauma Index: An aid in the evaluation of injury victims. J Trauma 1971;11:711–714.
31. Kaufmann CR, Maier RV, Rivara FP, et al. Evaluation of the Pediatric Trauma Score. JAMA 1990;263:69–72.
32. Bever DL, Veenker DL. An Illness-Injury Severity Index for nonphysician emergency medical personnel. EMT J 1979;3:45–49.
33. Gormican SP. CRAMS scale: Field triage of trauma victims. Ann Emerg Med 1982;11:132–135.
34. Baxt WG, Jones G, Fortlage D. The Trauma Triage Rule: A new, resource-based approach to the prehospital identification of major trauma victims. Ann Emerg Med 1990;19:1404–1406.
35. Clemmer TP, Orme JF, Thomas FT, et al. Prospective evaluation of the CRAMS scale for triaging major trauma. J Trauma 1985;25:188–191.
36. Ornato J, Mlinek EJ, Craren EJ, et al. Ineffectiveness of the trauma score and the CRAMS scale for accurately triaging patients to trauma centers. Ann Emerg Med 1985;14:1061–1064.
37. Levy PS, Goldberg J, Rothrock J. The Revised Estimated Survival Probability Index. Public Health Rep 1982;97:452–459.
38. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: The TRISS method. J Trauma 1987;27:370–378.
39. Cales RH. Injury severity determination: Requirements, approaches, and applications. Ann Emerg Med 1986;15:1427–1433.
40. Hoyt DB, Hollingsworth-Fridlund P, Winchell RJ, et al. Analysis of recurrent process errors leading to provider-related complications on an organized trauma service: Directions for care improvement. J Trauma 1994;36:377–384.
41. Bensard DD, McIntyre RC, Moore EE, et al. A critical analysis of acutely injured children managed in an adult level 1 trauma center. J Pediatr Surg 1994;29:11–18.
42. Champion HR, Copes WS, Sacco WJ, et al. The major trauma outcome study: Establishing national norms for trauma care. J Trauma 1990;30:1356–1365.
43. Champion HR, Copes WS, Sacco WJ, et al. A new characterization of injury severity. J Trauma 1990;30:539–546.
44. Copes WS, Sacco WJ, Champion HR, et al. Progress in characterizing anatomic injury severity. In 33rd Annual Proceedings, Association for the Advancement of Automotive Medicine. Des Plaines, IL;1989.
45. Markle J, Cayten CG, Byrne DW, et al. Comparison between TRISS and ASCOT methods in controlling for injury severity. J Trauma 1992;33:326–332.
46. Hannan EL, Mendeloff J, Farrell LS, et al. Validation of TRISS and ASCOT using a non-MTOS trauma registry. J Trauma 1995;38:83–88.
47. Champion HR, Copes WS, Sacco WJ, et al. Improved predictions from ASCOT over TRISS: Results of an independent evaluation. J Trauma 1996;40:42–49.
48. Flora JD. A method for comparing survival of burn patients to a standard survival curve. J Trauma 1978;18:701–705.
49. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
50. Lett RR, Hanley JA, Smith JS. The comparison of injury severity instrument performance using likelihood ratio and ROC curve analysis. J Trauma 1995;38:142–148.
51. Rutledge R, Osler T, Emery S, et al. The end of the Injury Severity Score (ISS) and the Trauma and Injury Severity Score (TRISS). J Trauma 1998;44:41–49.
52. Hosmer DW, Lemeshow S. Applied logistic regression. New York: J Wiley and Sons; 1989:140–145.
53. Rutledge R, Fakhry S, Baker C, et al. Injury severity grading in trauma patients: A simplified technique based on ICD-9 coding. J Trauma 1993;35:497–507.
54. Rutledge R. Injury severity and probability of survival assessment in trauma patients using a predictive hierarchical network model derived from ICD-9 codes. J Trauma 1995;38:590–601.
55. MacKenzie EJ, Steinwachs DM, Shankar B. Classifying trauma severity based on hospital discharge diagnoses. Med Care 1989;27:412–422.
56. Osler T, Rutledge R, Deis J, et al. ICISS: An international classification of disease-9 based injury severity score. J Trauma 1996;41:380–388.
57. Knudson P, Frecceri CA, DeLateur SA. Improving the field triage of major trauma victims. J Trauma 1988;28:602–606.
58. Esposito TJ, Offner PJ, Jurkovich GJ, et al. Do prehospital trauma center triage criteria identify major trauma victims? Arch Surg 1995;130:171–176.
59. Committee on Trauma of the American College of Surgeons. Resources for optimal care of the injured patient. Chicago: American College of Surgeons; 1999.
60. Phillips JA, Buchman TG. Optimizing prehospital triage criteria for trauma team alerts. J Trauma 1993;34:127–132.
61. DeKeyser FG, Paratore A, Seneca RP, et al. Decreasing the cost of trauma care: A system of secondary inhospital triage. Ann Emerg Med 1994;23:841–844.
62. Cottington EM, Young JC, Shufflebarger CM, et al. The utility of physiological status, injury site, and injury mechanism in identifying patients with major trauma. J Trauma 1988;28:305–311.
63. Ochsner MG, Schmidt JA, Rozycki GS, et al. The evaluation of a two-tier trauma response system at a major trauma center: Is it cost effective and safe? J Trauma 1995;39:971–977.