Systematic Review of Trauma System Effectiveness Based on Registry Comparisons

Gregory J. Jurkovich, MD; Charles Mock, MD, PhD

From the Harborview Medical Center, Trauma Services, Seattle, Washington.


**Background:** Trauma registries offer distinct advantages and disadvantages when assessing the effectiveness of trauma systems. Detailed injury data and statistical comparisons that use TRISS methodology and the Major Trauma Outcome Study norms provide advantages over population-based or preventable death studies. However, miscodings and registry differences in injury severity coding limit the validity and generalizability of findings. The purpose of this study was to identify these strengths and weaknesses and to determine whether registry studies provide evidence of trauma system efficacy.

**Methods:** A systematic review of published literature assessing trauma systems effectiveness by using registry-based data.

**Results:** Eight of 11 articles reviewed provided comparable data and consistently demonstrated a 15 to 20% reduction in the risk of death comparing trauma system outcomes to Major Trauma Outcome Study norms.

**Conclusion:** These studies provide evidence of the effectiveness of trauma systems. However, future studies that use trauma registries would be strengthened by including both prehospital and postdischarge trauma deaths, standardizing trauma registry inclusion criteria and developing a contemporary national reference norm for trauma outcome.

**Key Words:** Trauma systems; Trauma registries; Evidence review; MTOS.

The purpose of this section is to review the evidence of trauma system effectiveness derived from studies that use trauma registry databases. Trauma registry data, as defined for purposes of this discussion, include data from patients who are brought to a hospital and survive long enough to be admitted. Current trauma registries may or may not include deaths in the emergency room and do not include deaths in the field or deaths occurring in patients shortly after hospital discharge. Usually, registries contain detailed information on the actual injuries, which makes adjustment for severity of injury possible. The registry data, thus, may be used for such comparisons as before versus after institution of changes in a trauma system; location 1 versus location 2; and one location versus national norms.

In this section, we propose to accomplish the following: review the methodology involved in the use of trauma registry data to evaluate the effectiveness of trauma systems; summarize the advantages...
and disadvantages of the trauma registry data in assessing trauma system effectiveness, especially in comparison to the other methods discussed at this conference; review 11 key studies that used registry-based data to evaluate trauma system effectiveness, with particular attention to methodology and conclusions; summarize the data present in the registry table and formulate recommendations for future studies and registry databases; briefly present an international perspective on trauma system design and performance.

METHODS

When discussing comparisons between groups of patients with any disease, a major issue to be addressed is whether any observed difference in outcome is due to differences in severity of illness. For trauma patients, this means there must be some way to compare the outcomes of groups of patients with different severity of injuries in a statistically valid manner.

Most of the trauma registry studies that address trauma system development use a similar method of analysis: Flora's $Z$-statistic for statistical comparisons; use of the Major Trauma Outcome Study (MTOS) as a national norm; and TRISS methodology for injury severity adjustment. As the analysis of the results for the 11 key studies relies heavily on these methods, a brief review is warranted.

**Flora's $Z$ -Statistic**

The statistical basis for most of the comparisons that use trauma registry data involves a comparison of proportions. Most readers are probably familiar with the use of the $\chi^2$ statistic for the comparison of proportions among two different groups: for example, the proportion of deaths in group A versus group B. Flora's $Z$-statistic involves a slightly different concept: the proportion of deaths among an observed group compared with a national or a population wide proportion. Mathematically, this so-called test statistic is equivalent to $Z = \frac{(A - E)}{SE}$, where $A = $ actual proportion of deaths; $E = $ expected proportion of deaths (e.g., population value); and $SE$ is the standard error of the observed value. This gives a so-called $Z$-score, which corresponds to a level of significance. That is, if the $Z$-score is greater than 1.96 (or more than 2 standard deviations from the mean), the two groups are considered to be different at the 0.05 significance level.\(^1\)\(^,\)\(^2\)

Simplistically, you could use these numbers to include all patients with any condition (e.g., not accounting for severity). However, a further refinement on this statistical comparison is to estimate the proportion of patients who would have died in the national norm or population, had they had the same disease severity as the patients in the observed group. Flora's original study looked at survival among burn patients and used the percentage burn as the determinants of probability of death, with population (or expected) mortality rates for each percentage body surface area burned coming from the National Burn Information Exchange database.\(^1\) Hence, the statistical analysis uses Flora's statistic: **Equation (1)** where $S$ is the sum of individual number of survivors, $\Sigma Pi$ is the sum of the probabilities of survival of the individual patients in the dataset, as calculated from the national norms, and this difference is divided by the standard error term, which is the square root of the sum of the products of individual probabilities of survival and individual probabilities of death. Mathematically, this is equivalent to the more simple formula shown earlier.

\[
Z = \frac{S - \Sigma Pi}{(\Sigma PiQ_i)^{0.5}} \quad \text{Equation 1}
\]
MTOS Database

The next step in understanding the trauma registry comparisons is the development of an equivalent national norm of mortality rates for trauma patients in general (i.e., not for burns only). This has been provided by the MTOS, a work initially organized by the American College of Surgeons Committee on Trauma. Data on trauma patients were pooled from multiple hospitals throughout the United States and Canada, with the stated purpose to "develop and test survival probability norms based on severity indices." The MTOS database was primarily designed to be used for quality assurance by hospitals and emergency medical services (EMS) organizations. Data collected on hospitalized patients included demographics, etiology, specific injuries (originally Abbreviated Injury Scale [AIS]-85 and Injury Severity Score [ISS] coding), and outcomes in the form of in-hospital mortality. The study ran from 1982 to 1989 and collected data on approximately 160,000 trauma patients admitted to 139 United States and Canadian hospitals, representing about 7% of the hospitals in North America at that time. A partial list of participating hospitals was published in 1988. In 1990, a publication that used the MTOS database provided regression coefficients used to calculate probability of survival based on the 80,544 patients entered between 1982 and 1987. Subsequent publications have revised the coefficients used in the TRISS regression analysis based on additional patients and AIS-90 coding for anatomic injury. However, most of the studies presented in this review used the TRISS coefficients present in 1990, which were based on AIS-85 coding.

A description of the MTOS database identifies its strengths and shortcomings. It contains 79% blunt injuries and 21% penetrating injuries, and reports an overall mortality rate of 9%. Importantly, about 85% of the patients had some type of objective measurement of injury severity: autopsy report, computed tomographic scan, other radiographs, or surgery. However, it is not clear how many patients had autopsy data. The MTOS database is not only patients from Level I trauma hospitals, but includes a mix of Level I-III hospitals, some "pediatric specialized hospitals," and a few "postdesignated" hospitals. It is also important to note that this is not a population-based registry in that it is not a random sample of all the hospitals that exist in North America caring for trauma patients at that time, and that data collection was voluntary.

The norms were established for adults (age, ≥15 years) for both penetrating and blunt injury mechanisms from data provided by 51 institutions including 15,754 blunt injury patients and 7,423 patients with penetrating injuries. There are no pediatric-specific norms. The regression weights for young adult (age, 15-55 years) patients with blunt injuries were used as a proxy for pediatric patients. Importantly, 11% of the data submitted was incomplete, and those patients were excluded. This exclusion could introduce a huge bias if those 11% excluded patients who were the most severely injured or who had the poorest outcome. Despite these limitations, the MTOS database remains the largest descriptive database available as a national norm and has been used in almost countless publications, trauma QA activities, and trauma system comparisons.

TRISS Scoring

With the availability of a national norm for trauma outcomes, and a statistical method to compare outcomes in one group with the national norm (Flora's statistic), comparisons of outcomes also needed a method to determine the probability of death in the observed group of patients. In Flora's original work, this was provided by the burn percentage (i.e., one variable). For general trauma patients in the MTOS, the outcome predictors are multifactorial and include the following: ISS score, denoting anatomic extent of injury; Revised Trauma Score (RTS), denoting the physiologic extent of injury; and Revised Trauma Score (RTS), denoting the physiologic extent of injury.
injury; age, reflecting comorbidity and the ability to respond to injury

Because all three variables are related to a dichotomous outcome (survival vs. mortality), a multivariate statistical technique, logistic regression, is used. The probability of survival of an individual patient can then be derived as follows: 

\[ P_s = \frac{1}{1 + e^{-b}} \]

where 

\[ b = b_0 + b_1(\text{RTS}) + b_2(\text{ISS}) + b_3(\text{A}) \]

A is a variable for age (0 for age < 55 years, 1 for ≥ 55 years). \( b_0 \) through \( b_3 \) are coefficients derived from the use of the entire MTOS database in the logistic regression analysis. This method allows a given individual patient in a data set (as from an individual hospital's trauma registry) to have an estimated probability of survival calculated. This is equivalent to estimating the percentage mortality among all patients with the same trauma score, same ISS, and same dichotomous age stratification, nationwide.

Such probabilities are calculated for every person in a trauma registry database. The Z-statistic is then calculated in the same manner as shown earlier. This Z-score gives a statistical probability, indicating the statistical likelihood that an individual hospital's survival rate is indeed different from the national norm. What this does not tell you is the magnitude of the difference. More detailed descriptions of the TRISS method are available.\(^6\)

**Effect Size and W-Score**

Magnitude of difference in outcome is also known as "effect size." This can be approached on either a multiplicative or additive scale. By using the multiplicative method, the rate of outcome (e.g., survival or mortality) in one group is compared as a ratio with the same rate in another group or in a national norm, to provide a relative risk. For example, if a given group had a mortality rate of 30%, compared with another group with 20%; the relative risk (RR) of death would be a factor of 1.5 (or 50%) higher in the first compared with the second group. With the additive method, the difference in the rates of outcome are subtracted to obtain a risk difference. For example, 30% minus 20% indicates a 10% difference in mortality.

Neither method is correct nor incorrect. Both portray the same data in slightly different ways. In most epidemiologic work, relative risks (or the mathematically similar odds ratios) are used. However, in most of the publications on trauma registry outcomes, it is the risk difference that has been used. The data are usually portrayed in differences in survival rates, which is the same as differences in mortality rates. Accompanying the Z-statistic, an indicator of statistical significance, is the W-statistic, an additive measure of effect size. This is usually expressed as 

\[ W = \frac{A - E}{N / 100} \]

That is, actual number of survivors (A) minus estimated number of survivors (E) divided by number of persons in the study (N), divided by 100 to obtain a percentage. This value can then be thought of as number of extra survivors (or less survivors) per 100 injured persons treated. One disadvantage of the W-statistic is that it is highly influenced by the denominator used in the study. For example, if all patients admitted to a trauma center are included in a study, the W-statistic is likely to be very low, even if there are meaningful differences between the groups being compared. This is so because many of the admitted patients will have very low injury severity and, hence, the number of deaths per 100 persons admitted will be small. However, if only patients with severe injuries are included in a comparison, the W-statistic is likely to be larger. Multiplicative measures of effect size are less likely to be so affected. More detailed descriptions of the W-statistic are available.\(^3\)\(^7\)

**TRAUMA REGISTRIES**

Trauma registries have become an essential component of trauma systems. But the information in a
registry depends in part on what each individual hospital (or system) deems as important information and the cost of obtaining these data. There are a number of commercially available databases for trauma registry use. Some institutions have developed their own trauma registry database. Hence, differences in trauma registry element content exist. Early registry use focused on simple data collection such as reporting how many gunshot wounds were admitted in the past year and allowed comparison to previous time frames. More mature registries allow comparison of trauma care within a region, such as allowing queries that illustrate the distribution of the most severely injured patients among presenting hospitals, thereby assessing field triage practices. As a result of these capabilities, the presence of a well-functioning registry has become an important component of the trauma designation and verification process. Nonetheless, 33% of ACS Verification Committee site visits found deficiencies in the hospital trauma registries. Thus, although considered important, there are clearly some faults and deficiencies in trauma registries. Trauma registries have some advantages in analyzing trauma outcomes. All trauma registries should have reasonably similar formats, making inter-registry comparisons easier. Studies based on registry data remove much of the subjectivity that is inherent in assessing outcome with preventable death studies. Hospital registries contain greater individual injury detail than population-based databases such as hospital discharge data collected by states, or the Fatality Analysis Reporting System of National Highway Safety Traffic Association. As a result, hospital registry data should allow better adjustment of risk of death by injury severity or other comorbid factors, as long as these data are recorded (accurately) in the registry.

There are some provisos and limitations to the use of registry-based data to compare outcomes. First of all, the data from the index center or center of interest should not be included in the reference database, or should make up less than 5% of the reference data set. Second, the number of survivors in a predictive death analysis should be at least five in both groups to have any statistical significance. Additionally, and perhaps most troublesome in registry data, is that there must be some method of controlling for extraneous factors that might otherwise account for the observed differences. Unfortunately, it is often impossible to know what those extraneous factors might be. An example of the extraneous factors influencing outcome in trauma patients but unrelated to injury severity are comorbid conditions such as cirrhosis or severe cardiac disease. There are nearly limitless other factors (e.g., water fluoridation, number of EMS units per population density, etc.) that may or may not influence outcome, and none of them are recorded in registry data.

The advantages of the use of trauma registry data to assess effectiveness of trauma care include: they generally contain detailed injury description and some physiologic information; most registries have a similar format for the data, which helps with comparisons; compared with preventable death studies, the statistical comparisons that use TRISS and the MTOS norms remove some of the subjectivity associated with assigning the probability of survival among observed deaths; compared with population-based data, there is more detail on injuries, which makes severity adjustment possible. The disadvantages of registry data include the following, with the first three being common to all large databases:

1. Missing data. For example, among patients in the MTOS database, 11% had insufficient information to calculate a TRISS score, primarily because of missing respiratory rates.
2. Misdodings. Most experienced trauma directors have had the experience of reviewing trauma deaths that have been detected by audit filters such as deaths with ISS less than 15. It is not infrequent to have many or even most of these be due to misdodings, e.g., the injuries were more severe than initially coded. There is no easy way for such problems to be detected when large datasets are used.
3. Differences in the manner in which coding is done at different hospitals, especially with regard
to ISS. This of course may greatly influence the interpretation of outcomes, by skewing the assessment of the underlying severity of the patients. This is especially an issue with regard to differing proportions of autopsies performed at different hospitals, because autopsy data usually increase the calculated ISS score.

4. The collection of data for national norms, the MTOS, was not population-based. That is, trauma centers were enrolled voluntarily and were not chosen randomly. Hence, the norms for the hospitals chosen may reflect an outcome different from that in all hospitals in the United States and Canada.

5. Compared with preventable deaths studies, there is less detail about cause of death, nor is there the consensus opinion as to whether the death was medically preventable.

6. Compared with population-based studies, registry data do not take into account prehospital deaths, for which a large amount of the action is taking place in terms of trauma system effectiveness.\textsuperscript{9-11}

**DETAILED EXAMINATION OF KEY REFERENCES**

1. Impact of a Trauma System on Outcome of Severely Injured Patients. Shackford et al., *Arch Surg.*, 1987.\textsuperscript{12}

   **Purpose:** This study compared outcomes of the well-organized San Diego trauma system with the MTOS norms.

   **Patients:** All trauma patients with a trauma score of less than or equal to 8, who were taken to any of the trauma centers in San Diego. There were 249 such patients. The study ran for 1 year, 1984 to 1985.

   **Methods:** TRISS, comparison with MTOS.

   **Findings:** Observed/predicted deaths = 182 of 209 (blunt and penetrating together). Z-statistic = 5.01 for blunt, 3.72 for penetrating ($p < 0.01$ for both). Effect size: RR of death versus MTOS = 0.87. \textit{W}-statistic = 10.8 more survivors for 100 patients treated (TS $\leq$ 8).

   **Conclusions:** The authors concluded that the improved survival (compared with national norms) was due to the improved organization of the San Diego trauma system.

   **Comment:** There are a number of issues in the selection of patients reviewed in this study that might significantly affect comparability to other studies. For example, it is unclear whether patients who were declared dead on arrival or whether those who died in the emergency department were included or excluded from this study. In addition, autopsies were done on all patients, with the information from the autopsy being used to calculate the ISS. This has the possibility of substantially increasing the ISS because of better determination of all injuries. This rigorous review of autopsy data may not have occurred in the comparison population represented by those centers contributing to the MTOS national comparison database. Finally, transfers to hospitals in the San Diego trauma system from other outside hospitals do not appear to be included and might be included in other databases, including MTOS.

2. Trauma Experience of a Rural Hospital. Zulick et al., *Arch Surg.*, 1991.\textsuperscript{13}
Purpose: To compare outcome at a single "rural" hospital to MTOS database patients.

Patients: 163 trauma patients, 98% with adequate data, cared for over 2-year time frame (1984 to 1985).

Methods: TRISS comparison with MTOS data used Z-scores and W-scores.

Findings: Fourteen predicted deaths, 12 observed deaths, Z-statistic = –0.9 (postsignificant). All deaths reviewed: two possibly preventable.

Conclusion: Outcome at a rural hospital was considered to be equivalent to Level I centers in MTOS databank.

Comment: This is an early example of the use of TRISS and MTOS databases to compare trauma outcomes in one hospital with national norms. Individual review of all deaths was possible, and ISS and Trauma Score possible on nearly every patient. An M-statistic of 0.95 suggests a good match of these two patient populations, adding further validity to the data. However, 40% of the patients were transfers, nearly all were blunt trauma, and 18% were pediatric patients, making this population distinctly different than the MTOS database. Furthermore, despite the authors contention, the MTOS database is not all from Level I trauma centers, and this "rural" hospital was functioning at what might be considered Level II status, but without burns or computed tomographic surgery.


Purpose: To compare injured patients in the Montreal area versus MTOS.

Patients: Severely injured, but salvageable (355 patients), with a complex sampling scheme to obtain representative samples from 26 Montreal area hospitals.

Methods: Comparison with MTOS. Logistic regression analysis that uses ISS and age but not trauma score.

Findings: Observed/predicted deaths = 70/39; Z-statistic = 6.7 (p < 0.01). Effect size: RR of death (standardized mortality ratio), 1.8. W = 8.7 extra deaths per hundred patients treated (severely injured).

Conclusions: Worsened survival compared with national norms, because of lack of organized trauma system.

Comment: Similar questions about patient selection can be raised regarding the methodology used in this article. Autopsy data were not used for ISS calculation. The lack of autopsy data may lead to a systematic "underestimation" of ISS in this trauma system, in comparison to the MTOS database, in which contributing institutions had varying rates of autopsies (see Discussion section of article 1). This is an example of possible ascertainment bias.

What about inclusion or exclusion of patients declared "dead on arrival" or dead in the emergency department? Most studies do not comment on the inclusion or exclusion of these deaths, although some studies do explicitly state that they exclude such patients from any outcome analysis. Any
center in which these deaths are included would be biased toward worsened outcome. Any center with an aggressive EMS system that brings in more patients in extremis (as opposed to having them die at the scene and not be included in hospital data) would be biased toward worsened outcome. This is a selection bias.

4. Improvement in Outcome from Trauma Center Care. Champion et al., *Arch Surg.*, 1992.\(^{15}\)

**Purpose:** To analyze the longitudinal effectiveness of trauma center care at one institution.

**Patients:** Single urban hospital, longitudinal study from 1977 to 1982. Patients were grouped into two sets: all blunt injury deaths and survivors with length of stay more than 1 day, all blunt injury deaths and survivors with ISS more than 15.

**Methods:** Yearly TRISS comparison (Z-scores and W-scores) to norms based on entire 6-year period of study.

**Findings:** A steady increase in Z- and W-scores was noted each year. Overall, there were 4.34 more survivors per 100 patients than predicted, and in the severely injured set, there were 13.45 more survivors per 100 patients over the 6 years. Lack of RTS data resulted in 35% loss in predictive power and twice the number of false negatives (predicted to live, but died).

**Conclusions:** Improved outcome from trauma was coexistent with growing institutional commitment and trauma system development.

**Comment:** This study demonstrated statistical and clinically significant changes in outcome at one institution over time. It demonstrates that intrainstitutional improvement is possible. However, it is unknown whether this improvement in care occurred simultaneously at other hospitals, and it is not proven that better outcome is in fact related to "trauma commitment." Missing data were problematic in that 22% of patients were missing some component of TRISS. Autopsy determined ISS score in all deaths may be different from other populations.

5. The Effect of Secondary Insults on Mortality and Long-Term Disability after Severe Head Injury in a Rural Region without a Trauma System. Wald et al., *J Trauma.*, 1993.\(^{16}\)

**Purpose:** To determine whether the absence of a trauma system adversely affect outcome after head trauma.

**Patients:** One hundred seventy consecutive patients with severe head trauma (Glasgow Coma Scale score <8; 12% of total head injuries) were admitted to the Medical Center of Vermont between 1980 and 1985.

**Methods:** One hundred fifty fields in the head trauma registry were used to compare patients with data from the National Trauma Coma Data Bank. Two groups were identified and stratified based on the presence or absence of hypoxia and hypotension. Outcome was determined by three-strata evaluation of the Glasgow Outcome Scale.

**Findings:** Nearly all blunt trauma. Similar demographics and clinical variables, pathology, and treatment were found between comparison groups. Mortality in those with either hypotension or hypoxia was twice that of controls. No significant difference in outcome was found between National...
Trauma Coma Data Bank and Vermont.

Conclusions: Hypoxia and hypotension adversely affect outcome, and the impact of these secondary insults is unaffected by the presence of a regional trauma system.

Comment: The methods used in this study are not comparable to other reports in this review. The incidence of hypotension reported in this study (26%) is lower than that reported in the reference group (35%) and much lower than reported in some other studies (up to 47%). There were also fewer motor vehicles crashes and more scene deaths, and there were no comparisons of associated injuries (i.e., ISS not compared). Perhaps most importantly, one could argue that Vermont has a de facto trauma system.


Purpose: To determine concordance between regional outcome and national trauma center norms in a state lacking a trauma system.

Patients: All trauma patients less than 15 years of age admitted to a New York state hospital in 1989.

Methods: Comparison of summary discharge data from NY State DOH records with data for National Pediatric Trauma Registry patients from 1985 to 1989. ICDM-9 codes were divided into nine anatomic categories, and MacKenzie ICD-9 map for ISS scoring was used. Primary outcome variable of interest was death.

Findings: There were more severe head, brain, skeletal, and internal organ injuries in the National Pediatric Trauma Registry databank. Crude mortality was fivefold higher in National Pediatric Trauma Registry from brain injury and twofold higher if ISS more than 25. When stratified for ISS categories, mortality was similar for most patients, but there remained a 10-fold higher mortality in NY state patients for those with moderately severe brain, skeletal, and internal injuries.

Conclusions: The lack of an organized pediatric trauma system is costing lives in NY state. National norms are essential for setting target survival goals.

Comment: This study is limited in several areas. There is no overall adjusted mortality comparison; ISS was stratified by categories, and not as a continuous variable; there is no reporting of scene or prehospital deaths; there is the likelihood of duplicate reporting of records for transfers; there are no comorbid adjustments; and there is the possibility of coding errors by using ICD-9-derived ISS scores as noted previously.


Purpose: To determine the incidence, severity, and outcomes of injury in a rural Level I center compared with MTOS data.

Methods: Retrospective registry comparison to MTOS data by using TRISS, Z-statistics, and M-statistics.

Findings: Predicted mortality, 138; observed mortality, 115. The Z-statistic was –3.30, and the M-statistic was 0.99. Neurotrauma was the primary cause of death. TRISS data were available on 83% of patients.

Conclusions: Rural trauma centers can have mortality results similar or better to MTOS, despite small numbers of patients.

Comment: Once again, these authors have made the common erroneous assumption that MTOS data are entirely from Level I trauma centers. Other limitations of the study include only 83% of the patients had adequate data for TRISS analysis; 10 of the observed 125 deaths (8%) did not have TRISS data; 21% of the deaths were excluded from analysis because they were declared dead on arrival; there is no autopsy data; the population is skewed to less severe injuries.

8. Analysis of a Rural Trauma Program Using the TRISS Methodology: A 3-Year Retrospective Study. Karsteadt et al., J Trauma., 1994.19

Purpose: To evaluate the effectiveness of a "model rural trauma project."

Patients: Two hundred sixty-six patients from the North Coast EMS Region of California with major trauma, defined as deaths, ISS more than 15, or trauma team activation.

Methods: Three-year retrospective study after implementation of model rural trauma project, involving five rural hospitals. ISS was hand calculated from AIS scoring. Data were collected on CDC Trauma Registry Database v. 2.0. TRISS was calculated and compared with MTOS survival norms.

Findings: Two hundred sixty-six patients identified from 1,017 total trauma patients. 5.2% of patients missing data. Overtriage, 0.9%; undertriage, 6.2%. Predicted mortality, 63; observed mortality, 54. Calculated Z-statistic, –2.33, p = 0.020 (less mortality than expected); M-statistic, 0.66 (more severe injuries than MTOS).

Conclusions: Improved survival compared with MTOS database norms.

Comment: The model rural trauma project program was designed to avoid ACS center designation and commitment of resources. Survival was better than expected, despite limiting the analysis to only the most severely injured patients. Six of the unexpected deaths were attributed to comorbidity or prolonged prehospital care, emphasizing the importance of these two confounders that may or may not be captured in registry data, and are infrequently controlled for in outcome analyses. These findings are tempered by the previously cited limitations associated with the use of MTOS as a criterion standard. Overall, this study emphasized the importance of complete EMS system improvement not just hospital or surgeon resources.

9. The Early Effects of Implementing American College of Surgeons Level II Criteria on Transfer and Survival Rates at a Rurally Based Community Hospital. Norwood et al., J Trauma., 1995.20

Purpose: To determine the effects of ACS Level II implementation on transfers and survival.
Patients: All trauma admission to one hospital in time period A (13 months before ACS Level II status) compared with time period B (14 months after), 1992 to 1993.

Methods: A "before and after" study at one hospital. AIS was used to calculate ISS. MTOS comparison was made at each time period.

Findings: Time period A: 699 total trauma patients, 54 deaths, 189 patients with ISS more than 15, 189 transfers. Time period B: 862 total trauma patients, 69 deaths, 297 patients with ISS more than 15, 297 transfers. There were more transfers and more severely injured patients seen after implementation of trauma center status. Although overall survival was not significantly different from MTOS norms during either time period, patients with ISS more than 15 had survival improvement from 71.4 to 76.8% (W = 5.4, RR of death = 0.81, not significant), and those patients least likely to survive (p ≤ 25%) had improved outcome from 7.5% survival before to 25.5% survival after trauma center implementation (W = 18, RR of death = 0.81, p = 0.3).

Conclusions: Implementing ACS trauma center guidelines have the affect of increasing transfer of seriously injured patients and improving survival in the most critically injured patients.

Comment: Although this study supports the contention that ACS Level II trauma center resources improves outcome in the most severely injured patients, there remains the concern that patients who are transferred are somehow different from those who are primarily triaged from the field to an institution. A high percentage of transfer patients in any population may represent a form of selection bias, and ISS or RTS may be inadequate injury descriptors in comparing a transfer population to primary triage populations.

10. An Evaluation of Patient Outcomes before and after Trauma Center Designation Using Trauma and Injury Severity Score Analysis. Stewart et al., J Trauma., 1995.21

Purpose: A comparison of injury outcome in the main trauma hospital in southwest Ontario, Canada, before and after trauma center designation. Designation was accompanied by increased governmental commitment and increased infrastructure.

Patients: Patients admitted to Victoria Hospital (London, Ontario) with ISS more than 12 from motor vehicle crashes (n = 303). Two 12-month periods were compared (1989/1990 vs. 1992/1993).

Methods: TRISS, comparison with MTOS: done separately for each time period.

Findings: Before: Z = 1.34. No statistical difference from MTOS. Twelve unexpected deaths. Because the Z-statistic was not significant, no W-statistic was calculated. After: Z = 2.97 (p < 0.05 vs. MTOS). Six unexpected deaths. W = 5.6 versus MTOS.

Conclusions: Trauma center designation and improved infrastructure results in improved outcome.

Comment: Autopsies were used for ISS calculations. However, the percentage of injured trauma patients who underwent autopsy was not mentioned. Interhospital transfers were included in this study, and the proportion of transferred patients increased from 61% before designation to 74% after designation. There also is the possibility that changes in aspects of health care other than trauma center designation occurred between these time periods and may have accounted for the observed difference. An alternative analysis would have been to compare mortality rates before and after changes, adjusted by TS, ISS, age, and mechanism (TRISS methodology), without reference to
MTOS. In the current method, we do not really know whether the improvement in survival at their facility was statistically significant nor do we know the magnitude of the improvement.


Purpose: To develop new TRISS coefficients that use data from patients within a single region (Ontario, Canada) to provide a better tool with which to assess trauma hospital performance in a region.

Patients: Three years of data on 5,258 patients (ISS > 12) presenting to 11 hospitals in the Providence of Ontario.

Methods: The data were subjected to a traditional TRISS-like methodology that used MTOS coefficients. These same data were then used to generate "region-specific" coefficients and, by using the subsequent years data, a similar analysis was used.

Findings: The use of traditional MTOS coefficients produced an expected mortality of 21.2% (1,115.6 of 5258) with a $Z$-statistic for the entire providence of –14.102. Based on the validation set, the "regionalized" coefficients generated an overall $Z$-statistic of –2.540.

Conclusions: The authors conclude that new coefficients were more accurate in predicting patient outcomes. Although this article is of a different format than the others reviewed for this section, it is an important example of the limitation of using the MTOS database as the national or international reference norm. Few Canadian hospitals were included in MTOS, a database that is now over a decade old. This study suggests that overall improvement in care over that time span has occurred. When compared with MTOS data, nearly every Ontario hospital did well in 1994 to 1995. However, when more contemporary and region-specific coefficients were derived, a different (more accurate?) picture of performance is observed. This study also demonstrated the limitation of MTOS in excluding those patients in whom RTS cannot be calculated because of intubation precluding Glasgow Coma Scale score determination. By using a modification of TRISS methodology proposed by Offner et al., such patients were included in developing the new coefficients for TRISS calculations of survival probability.23 The article also suggests that regional specific analyses may identify outliers that would be overlooked by national or international norms.

SUMMARY OF KEY TRAUMA REGISTRY STUDIES

Table 1 is a compilation of 11 significant studies published between 1987 and 1997 that used trauma registry data to evaluate the effectiveness of trauma systems. Although there are likely many other articles in this same vein, we believe these represent key analyses. In Table 2, we have tried to simplify and standardize this analysis by looking solely at the change in the relative risk of death that occurred with implementation of a trauma system or program. We also give some measure of the significance or impact of this change with the W-statistic, i.e., the additional or unexpected number of lives saved per 100 patients treated. We further subgroup these studies into three categories: those that evaluate an existing trauma system/program in comparison to the MTOS database, those that had no trauma program compared with MTOS, and those programs that use a temporal analysis to compare themselves with MTOS or conducting a before and after implementation of a trauma program/system. Note that not all 11 studies are amenable to this standardized analysis, but the 8 that are provide a consistent finding.
The striking feature of the four studies comparing in-place trauma system with the MTOS database is that they all have essentially the same outcome, namely a relative risk of death that is about 0.85, meaning that these programs have a 15% decrease in their death rate compared with MTOS database. These four studies are from divergent programs, including rural locations, a developing trauma system, and one that is very mature. Despite this diversity, there is a striking similarity in trauma center designation, resource commitment, and decreased mortality in these hospitals compared with a national reference standard.

There is only one article comparing a region (Montreal) that is lacking any trauma system to a reference database (MTOS). In this single report, a region lacking any trauma system is shown to have worse, with an 80% higher relative risk of death, and nearly nine more deaths per 100 compared with the MTOS database. It is likely that there is a publication selection bias in that editors do not like to publish negative result articles, and authors do not like to submit reports documenting adverse outcomes or poor results, hence, the paucity of articles that show increased deaths in areas lacking a trauma system. It should be noted that Sampalis and colleagues have since published improved survival in Quebec with regionalization of trauma care.

The third group of articles summarized in Table 2 are those that compare trauma mortality in the same hospital or region before implementation of any changes in trauma care and after such changes have been made. The article by Champion et al. is an example. In this article, patients were divided into two subgroups: all patients who were admitted, and the most severely injured patients as judged by ISS more than 15. Although a RR of death could not be calculated from the data presented, the W-statistic for the most severely injured patients improved from −8.2 to +1.7 over the 6 years of increasing commitment and experience. This is a consistent finding in other studies, in that the most significant impact of a trauma system is in the care of the most severely injured patients. Two other studies noted in the table also use a "before and after" comparison method, but rather than an internal comparison, they use the MTOS as a reference standard for both time periods. The relative changes demonstrated are consistent, with the article by Norwood et al. demonstrating a 0.81 RR of death after trauma system implementation, and the article by Stewart et al. showing an improved W-statistic of 5.6 after trauma system implementation.

What can we conclude from this overview of trauma registry data? We believe these studies provide evidence that trauma systems are effective. We struggle with a descriptor of the evidence. Is it good evidence? Is it great evidence? Is it acceptable evidence? Is it marginal evidence? What it is not, is
prospective randomized clinical trial evidence, sometimes referred to as "Class I" evidence. It cannot be. Trauma registries are not designed that way, and they will not provide that kind of data. Hence, there are problems with the data. There clearly is a publication bias. Publications of negative results or poor outcomes are rare. Most studies suffer from a time lag between either the reference standard or internal before/after comparisons. Contemporary side-by-side comparisons of one trauma system versus no trauma system are also unusual, with the Oregon-Washington study of Mullins et al. a rare example. Also troubling is the fact that the reference standard used for these studies has been static, with data submission to the MTOS database ending in 1989. It is likely that comparing current trauma outcomes to over 10-year-old trauma registry data would nearly always show improvement. The reference standard is also not population-based, but neither is it limited to only Level I trauma centers. There are also questions about variable inclusion criteria, and of course, individual patient review is missing from these registry databases. Finally, there is significant variability in describing injury severity based on the method used to calculate ISS, a problem perhaps particularly troublesome in penetrating injury populations.

RECOMMENDATIONS FOR IMPROVED TRAUMA REGISTRY ANALYSES

Although the evidence of trauma system effectiveness provided by registry-based data is fairly consistent and uniform, the above discussion highlights areas in need of improvement. Trauma registries are powerful tools, well established in trauma centers, and a useful resource for evaluating trauma care. However, they can be improved. Our recommendations are as follows: A contemporary national reference norm for trauma outcome is needed. The MTOS database is not contemporary, and it is not population-based. A national reference database should be stratified by level of commitment of the trauma center. One ought to be able to compare outcomes at a given hospital with the national norms of Level I trauma care. Ideally, a population-based registry should allow comparisons of a given hospital to the outcome of trauma patients treated at a representative sample of all the hospitals in the United States, whether designated trauma centers or not.

Contemporary national norms should include both prehospital and postdischarge mortality rates. Without these data, there is limited ability to evaluate the efficacy of a "system;" rather, one is simply evaluating hospital care alone.

Uniform and standardized trauma registry inclusion criteria and a consistent method of injury scoring are required of all institutions. Variability in ISS scoring methods, autopsy data inclusion, and patient inclusion criteria have further limited the usefulness of registry databases. Comorbid factors known to significantly affect outcome should be included in all registries. Functional outcomes need to be assessed in trauma patients. Simple inhospital mortality is an inadequate measure of trauma system effectiveness. A uniform tool or method of assessing functional outcome in trauma patients needs to be developed, tested, and validated. Results of such measurements need to be added to trauma registry databases.

INTERNATIONAL PERSPECTIVE

We will conclude this section with the results of a study comparing the outcomes of trauma systems in three different nations, findings we feel have direct bearing on this symposium. In a recent study, we compared the mortality for all seriously injured persons (defined as ISS > 9) in three cities: Seattle, Washington; Monterrey, Mexico; and Kumasi, Ghana. These locations differ greatly in their economic status, in that the per capita gross national product of the three countries is $23,000, $3,000, and $310, respectively. They also differ greatly in their trauma treatment capabilities. Seattle
has one of the most advanced EMS systems in the world. Monterrey has a basic EMS system. Kumasi has no EMS system, with injured persons brought to the hospital by relatives and public transportation. There are major differences in the hospital capabilities in each location as well.

Figure 1 displays the mortality pattern in these regions, with site of death determined to be either prehospital, emergency department, or elsewhere in the hospital. Overall survival among these cohorts of seriously injured patients increased with increasing economic status in the three cities: Kumasi (36%), Monterrey (45%), and Seattle (65%). Figure 1 shows that improvement in survival primarily is due to a decrease in prehospital deaths. The percentage of all seriously injured persons dying in the field declined progressively: Kumasi (51%), Monterrey (40%), and Seattle (21%). The second major site of deaths was in the emergency department. In Monterrey, 11% of deaths of seriously injured persons occurring at that point in the trauma system.

These data have three implications regarding trauma system efficacy. First of all, mortality patterns of one of the best trauma systems in the world compared with one of the least advanced countries demonstrates that the overall relative risk of death is twice as high as in the least advanced area. Thus, when evaluating trauma systems within the United States, we are looking at a much more homogeneous situation and the differences should be less and more difficult to discern. Second, these data show the great importance of evaluating and understanding the prehospital phase of trauma care in any trauma system. Finally, in this conference, we are restricting ourselves to discussion of trauma systems in North America. Any conclusions that we reach may or may not have pertinence to other environments. However, the essentials of this discussion and symposium can be summed up in a question: "Can we improve outcome after trauma by better organization of existing facilities and infrastructure, with minimal additional resource input?" We believe that the answer to this question is "Yes," and that better organization can improve outcomes in resource-poor environments and in our own country.

REFERENCES


Address for reprints: Gregory J. Jurkovich, MD, Trauma Services, Harborview Medical Center, 325 Ninth Avenue, Box 359796, Seattle, WA 98104; email: jerryj@u.washington.edu