

In Search of Benchmarking for Mortality Following Multiple Trauma: A Swiss Trauma Center Experience

Ida Füglistaler-Montali · Corinna Attenberger ·
Philipp Füglistaler · Augustinus L. Jacob ·
Felix Amsler · Thomas Gross

Published online: 20 August 2009
© Société Internationale de Chirurgie 2009

Abstract

Background The manifestations associated with non-survival after multiple trauma may vary importantly between countries and institutions. The aim of the present study was to assess the quality of performance by comparing actual mortality rates to the literature.

Methods The study involved evaluation of a prospective consecutive multiple trauma cohort (injury severity score, ISS > 16) primarily admitted to a university hospital. Univariate and multivariate testing of routine parameters and scores, such as the Trauma and Injury Severity Score

(TRISS), was used to determine their predictive powers for mortality.

Results The 30-day mortality of 22.8% ($n = 54$) exactly matched predicted TRISS versions of Champion or the Major Trauma Outcome Study for our 237 multiple trauma patients (42.8 ± 20.9 years; ISS 29.5 ± 11.5). Univariate analysis revealed significant differences between survivors and non-survivors when compared for age, ISS, Glasgow coma scale (GCS), pulse oximeter saturation (SapO₂), hemoglobin, prothrombin time, and lactate. In multivariate analysis, age, ISS, and GCS ($P < 0.001$ each) functioned as major independent prognostic parameters of both 24 h and 30-day mortality. Various TRISS versions hardly differed in their precision (area under the curve [AUC] 0.83–0.84), but they did differ considerably in their level of requirement, with the TRISS using newer National Trauma Data Bank coefficients (NTDB-TRISS) offering the highest target benchmark (predicted mortality 13%, Z value -5.7) in the prediction of 30-day mortality.

Conclusions Because of the current lack of a single, internationally accepted scoring system for the prediction of mortality after multiple trauma, the comparison of outcomes between medical centers remains unreliable. To achieve effective quality control, a practical benchmarking model, such as the TRISS-NTDB, should be used worldwide.

I. Füglistaler-Montali · C. Attenberger · P. Füglistaler ·
A. L. Jacob · F. Amsler · T. Gross
Computer Assisted Radiology & Surgery Switzerland
(CARCAS), University Hospital Basel, Spitalstrasse 21,
4031 Basel, Switzerland

F. Amsler
www.amslerconsulting.ch

I. Füglistaler-Montali
General and Visceral Surgery, University Hospital Münster,
Waldeyerstr. 1, 48149 Münster, Germany

C. Attenberger
General and Visceral Surgery, Klinikum Augsburg, Steglinstr. 2,
86156 Augsburg, Germany

P. Füglistaler
Vascular Surgery, St. Franziskus Hospital, Hohenzollernring 72,
48145 Münster, Germany

A. L. Jacob
Department of Interventional Radiology, University Hospital
Basel, Petersgraben, 4031 Basel, Switzerland

T. Gross (✉)
Realpstr. 57, 4057 Basel, Switzerland
e-mail: tgross@uhbs.ch

Introduction

The quality of emergency care and outcomes after multiple trauma may vary importantly between different countries and individual hospitals [1]. Survival of the multiply injured patient remains the primary objective of treatment, but mortality continues to be high, especially in the presence of major head trauma or serious hemorrhage [2]. The

probability of survival depends on the trauma, the patient, and the quality of care received. There is an increasing interest in risk scoring systems in surgery and external benchmarking for trauma center performance [3–5]. Trauma scores may serve as instruments of quality control for the systematic comparison of patients and institutions [6]. Even though outcome prediction was found to be insufficient for individual decision making in clinical situations [7], it might serve as a tool for quality assurance in the comparison of therapeutic results. To improve the quality of care, it should be possible to compare the performance data from different institutions based on an internationally accepted standard [8]. Susan Baker's statement on this issue should be remembered: "If you have never felt the need for any type of severity scoring system, then you probably have never had to explain how it is that the survival rate of 58% in your trauma center is actually better than the survival rate of 97% in some other hospital where the patients are much less seriously injured" [9]. The multifactorial sequence of actions following multiple trauma and the various possible confounders make it obvious that one single parameter cannot be sufficient for the comparison of data, but scores that combine several variables may be helpful. Thirty years ago original investigations reported on trauma severity scoring to predict mortality and evaluate trauma care [10, 11]. Over the years, several modifications of original scores have been developed, and more complex prediction models have been designed to improve comparison of a hospital's expected outcomes with its actual outcomes [12]. However, the fact remains that there is still no international consensus on which score is the most reliable. The relevant articles are numerous, and reports are often confusing because of conflicting findings or noncomparable data. Any interpretation of raw hospital data has to take possible confounders into account—such as different trauma populations or health care structures—with the consequent need for adjustment in the case mix being studied [13–15]. But studies in the field continue to publish raw data—for example, mean mortality rate—without including sufficient information on the composition or management of the cohort investigated [16]. This makes any reasonable comparison of the quality of treatment between different centers impossible. In certain areas, participation in a regional multicenter database program, such as the National Trauma Data Bank [17], the Trauma Audit Research Network database [18], or the German Trauma Registry [19], may offer an opportunity to compare data from the contributing hospitals. Even so, data recruitment for national registries varies widely, a finding that underlines the lack of uniform benchmarking [20]. Furthermore, there are still many regions in the world without any form of regional data collection. Irrespective of the extensive body of literature on the subject, an internationally accepted standard of mortality prediction to permit adequate quality assessment is still missing [5, 8, 21, 22].

Our investigation posed two main questions: How good is the management of multiply injured patients in this institution in terms of actual mortality rate? and Which prediction model will emerge as the most useful for routine clinical use? In the context of a typical single-center scenario, we therefore tested the practicability of several scoring systems as discussed in the literature and composed of routine parameters used worldwide in the emergency treatment of multiple trauma patients. We focused mainly on comparison of the multiple variations reported for the most frequently applied mortality prediction model, the Trauma and Injury Severity Score (TRISS) [11, 23, 24].

Materials and methods

Patients and procedure

From August 2001 to April 2005, 506 consecutive trauma patients arrived in the emergency room (ER) of our university hospital with suspected multiple trauma and were entered into a prospective database. Multiple trauma emergency treatment guidelines were followed according to international standards: ER diagnostics routinely included Focused Assessment with Sonography for Trauma (FAST) [25] and conventional x-ray for thorax and pelvis, followed by multislice CT, if needed (including clarification of the cervical spine). In every case, at least the trauma team leader (trauma surgeon or general surgeon), but more often several members of the interdisciplinary resuscitation team, had completed an Advanced Trauma Life Support (ATLS) course [26]. Multiple trauma patients were defined as trauma victims where (1) at least two Abbreviated Injury Severity (AIS) [27] regions were affected, and (2) the Injury Severity Score (ISS) was >16 [28]. Patients with mono-trauma, $ISS \leq 16$, and those secondarily admitted from another hospital were excluded from the analysis ($n = 269$; 30-day mortality rate 8.6%). For every emergency case medical students specifically trained for the study but not involved in the treatment of patients were alerted and called to join the team on duty. They were present during the diagnostic and therapeutic process in the resuscitation bay, and they documented the clinical procedure as well as all available laboratory data in a standardized fashion. Pre-hospital variables were extracted from the ambulance or helicopter documentation. Patients' demographic data and the variables needed to determine all scores were collected prospectively on admission, and the first available value (prehospital or emergency room) was used for analysis: age, ISS [27], GCS [29]; systolic blood pressure [SBP] [30]; arterial hemoglobin saturation by pulse oximetry [SapO₂] [31]; breath rate [BR] [32]; heart rate [HR] [33]; shock index [SI] [34]; hemoglobin [Hb] [35]; prothrombin time

test [Quick's value, PT] [19]; base excess [BASE] [36]; and lactate [37]. The whole data collection was subsequently reviewed and completed by fellow surgeons. The calculation of scores and models was undertaken according to the cited literature for the AIS and ISS, respectively, based on the follow-up data obtained on hospital discharge [27, 38, 39]: ISS, Revised Trauma score [RTS] [40]; Triage Revised Trauma score [T-RTS] [40], PTS [41]; Bouamra score [42], Simplified Acute Physiology Score [SAPS II] and resulting predicted mortality [43], and the TRISS in the versions of Boyd (Boyd-TRISS) [11], Champion (Champion-TRISS) [23], the Major Trauma Outcome Study (MTOS-TRISS; www.trauma.org/index.php/main/article/387), the Trauma Audit & Research Network (TARN-TRISS) [44], and the National Trauma Data Base (NTDB-TRISS) [24].

Death and time of death were recorded for every patient, and mortality rate was assessed after 24 h, 30 days, and 2 years following trauma. The study was approved by the local ethical committee, and informed consent was obtained from patients or relatives if possible.

Statistics

The results are presented as mean \pm standard deviation (SD), if not stated otherwise. All statistical tests were two-tailed. Student's *t*-test was used for comparison of means in normally distributed data of continuous variables; analysis of variance (ANOVA) was used for similar criteria in three or more unpaired subsamples. Chi-square analysis was used to test categorical data. Routinely documented variables suspected or known from the literature to be possible factors associated with mortality were analyzed, resulting in the parameters for univariate analysis with the primary endpoint mortality (24 h and 30 day): age, ISS, GCS, SBP, SapO₂, BR, HR, SI, Hb, PT, BASE, lactate, RTS, PTS, and single TRISS-versions.

Stepwise regression analysis was performed by including all factors that were found to be significant in univariate analysis. Starting with the major variables age, ISS, and GCS in the first three steps, other significant factors of the univariate analysis were added separately in a fourth step, whereby missing data were replaced by the mean. The stepwise model was executed in order to evaluate any additional information obtained from subsequently added parameters using mortality as the dependent variable and any variable under investigation as an independent variable. Nagelkerke R² was calculated to estimate variance of the model, and the chi-square statistic was used to calculate the significance of the improvement of the model.

Mortality scores were calculated in accordance with the relevant descriptions in the literature. Most of these scores contain the RTS, which is calculated using GCS, blood pressure, and breath rate as a subscore. As the RTS is

reported to have a high rate of missing values in the literature, scores were (1) calculated and compared for the sample without missing values ($n = 144$) and (2) coded for missing breath rates (90 of 237 cases) according to the nonpathological category of breath rate to avoid overestimation of mortality, and missing blood pressures were entered as the mean of the fully documented sample (3 of 237 cases) [44]. Calculation of TRISS values for the group of patients with missing variables (breath rate) and comparison with those for the rest of the study cohort did not reveal any significant difference between groups in terms of age, ISS, GCS, predicted SAPS II, or observed 30-day mortality. For example, initial breath rate was missing in 93 patients (mean age: 43.4 ± 20.2 ; mean ISS: 29.3 ± 10.9 ; mean GCS: 10.9 ± 4.4 ; predicted SAPS II mortality 76.9%; 30-day survival 73.1%) compared with 143 patients with these data (mean age: 42.5 ± 21.4 , $P = 0.732$; mean ISS: 29.6 ± 11.9 , $P = 0.812$; mean GCS: 10.2 ± 4.7 , $P = 0.239$; predicted SAPS II mortality 75.0%, $P = 0.598$; 30-day survival 80.9%, $P = 0.229$).

Scores were tested by generation of receiver operating characteristic curves (ROCs) and by comparison of the areas under the curve (AUC), which are reported with 95% confidence intervals (CI). The precision of the models investigated (AUC) differed only minimally if patients with missing values were excluded from the subsequent analysis or if missing values were substituted in this standardized fashion. Data are only shown for the latter.

The *Z*-statistic was calculated for definitive outcome-based evaluation (DEF) [45]. In DEF, Flora's *Z*-score quantifies the difference between the actual number of deaths (or survivors) in the test subset and the predicted number of deaths (or survivors). The formula for calculating *Z* is: $Z = D - (q_i/p_iQ_i)$, where *D* is the actual number of deaths, *Q_i* is the predicted probability of death for a patient *i*, *q_i* the predicted number of deaths and *p_i* the predicted *P_s* for patient *I* [46]. A *Z*-score with an absolute value of >1.96 is statistically significant [47]. As our study population (due to the inclusion criteria presenting with a higher percentage of severely injured patients; $M < 0.88$) differed importantly from the Major Trauma Outcome Study cohort, further statistical analysis or interpretation was not undertaken [47, 48].

To calculate sensitivity and specificity statistics, mortality scores were dichotomized around the value of 0.5. Data were analyzed with SPSS version 13.01 (SPSS Inc., Chicago, IL), and a *P* value <0.05 was considered significant.

Results

The prospective cohort consisted of 237 consecutive multiply injured patients (mean age: 42.8 years; 73.4% male;

3.4% penetrating trauma) with a mean ISS of 29.5 (Table 1). Trauma cases were the consequence of traffic accidents (65.0%), falls from a height (18.6%), or other causes (16.4.9%). The average length of stay in the intensive care unit (ICU) was 6.2 days and at our hospital 13.0 days. The 24-h mortality rate was 10.1% ($n = 24$); 30-day mortality, 22.8% ($n = 54$); overall 2-year mortality, 24.1% ($n = 57$).

Univariate analysis on 30-day mortality found significantly higher values—that is, worse values—for non-survivors for age, GCS, SpO₂, Hb, PT, lactate, hospital stay, ISS,

T-RTS, Champion-TRISS, and NTDB-TRISS, as well as SAPS II, in comparison to survivors (Table 2). None of the variables under investigation showed an obvious cut-off value beneath which survival could be ruled out definitively.

The prediction model for 30-day-mortality including the most significant variables age, ISS and GCS showed an overall variance of 46.7% (Table 3). Further stepwise logistic regression found that both the initial hemoglobin and the prothrombin values still added some significant information to the model (2.1%, $P = 0.02$ resp. 2.4%, $P = 0.32$). When both variables (Pearson correlation 0.50)

Table 1 Descriptive values and predicted survival rate for patients in series

| | <i>N</i> | Mean | Standard deviation | Median | Minimum | Maximum |
|-------------------------|----------|--------|--------------------|--------|---------|---------|
| Age, years | 237 | 42.84 | 20.92 | 39.3 | 7.3 | 91.8 |
| GCS | 237 | 10.44 | 4.60 | 13.0 | 3.0 | 15.0 |
| SapO ₂ , % | 229 | 0.95 | 0.07 | 1.0 | 0.6 | 1.0 |
| SBP, mmHg | 234 | 128.21 | 30.06 | 130.0 | 0.0 | 220.0 |
| HR, beats/min | 228 | 94.07 | 24.04 | 90.5 | 37.0 | 160.0 |
| SI | 225 | 0.78 | 0.31 | 0.7 | 0.3 | 2.2 |
| BR, respirations/min | 147 | 16.92 | 7.80 | 15.0 | 1.0 | 40.0 |
| Hb, g/d | 213 | 11.84 | 2.85 | 12.1 | 3.0 | 19.0 |
| PT, % | 197 | 0.75 | 0.25 | 0.8 | 0.1 | 1.2 |
| BASE, mmol/l | 113 | -3.07 | 5.50 | -4.0 | -15.0 | 12.0 |
| Lactate, mmol/l | 81 | 3.50 | 3.19 | 2.0 | -0.5 | 18.0 |
| AIS I | 237 | 2.84 | 1.65 | 3.0 | 0.0 | 6.0 |
| AIS II | 237 | 0.78 | 1.01 | 0.0 | 0.0 | 4.0 |
| AIS III | 237 | 2.58 | 1.60 | 3.0 | 0.0 | 5.0 |
| AIS IV | 237 | 1.11 | 1.38 | 0.0 | 0.0 | 5.0 |
| AIS V | 237 | 1.94 | 1.40 | 2.0 | 0.0 | 5.0 |
| AIS VI | 237 | 0.48 | 0.58 | 0.0 | 0.0 | 3.0 |
| ISS | 237 | 29.48 | 11.53 | 27.0 | 17.0 | 75.0 |
| T-RTS | 237 | 10.45 | 1.86 | 11.0 | 3.0 | 12.0 |
| RTS | 237 | 6.38 | 1.56 | 6.9 | 1.8 | 7.8 |
| PTS | 237 | 18.55 | 12.94 | 17.0 | 0.0 | 71.0 |
| SAPS-II | 237 | 34.54 | 19.21 | 31.0 | 3.0 | 93.0 |
| ICU stay, days | 211 | 6.15 | 9.63 | 3.0 | 1 | 104 |
| Hospital stay, days | 237 | 13.00 | 12.96 | 11.0 | 1 | 105 |
| Boyd- TRISS survival | 237 | 0.79 | 0.28 | 0.93 | 0.00 | 0.99 |
| Champion-TRISS survival | 237 | 0.78 | 0.29 | 0.93 | 0.00 | 0.99 |
| MTOS-TRISS survival | 237 | 0.77 | 0.28 | 0.91 | 0.00 | 0.99 |
| NTDB-TRISS survival | 237 | 0.87 | 0.19 | 0.94 | 0.00 | 0.99 |
| TARN-TRISS survival | 237 | 0.72 | 0.31 | 0.85 | 0.00 | 0.99 |
| Bouamra score survival | 237 | 0.69 | 0.32 | 0.82 | 0.01 | 0.99 |
| BMTPM survival | 237 | 0.77 | 0.26 | 0.88 | 0.00 | 0.99 |
| SAPS II survival | 237 | 0.76 | 0.27 | 0.89 | 0.03 | 0.99 |

GCS Glasgow Coma Scale, SapO₂ saturation by pulse oximetry, SBP systolic blood pressure, HR heart rate, SI shock index, BR breath rate, Hb hemoglobin, PT prothrombin time test, BASE base excess, lactate, AIS Abbreviated Injury Scale, ISS injury severity score, T-RTS Triage Revised Trauma score, RTS Revised Trauma score, PTS Polytrauma score, SAP II Simplified Acute Physiology score, ICU intensive care unit, TRISS Trauma and Injury Severity score, in the versions of Boyd (*Boyd-TRISS*) or Champion (*Champion-TRISS*), or the Major Trauma Outcome Study (*MTOS*), or the Trauma Audit & Research Network (*TARN*), or the National Trauma Data Base (*NTDB*) as *NTDB-TRISS*, the Bouamra score, *BMTPM* Basel Multiple Trauma Prediction of Mortality

Table 2 Univariate analysis for 30-day survivors versus non-survivors

| Variable | 30-day-mortality | | N non-survivors | Mean \pm SD or % survivors | Mean \pm SD or % non-survivors | P value | Mean diff. | df | t | CI 95% of the difference | |
|----------------------|------------------|-------|-----------------|---------------------------------|----------------------------------|---------|------------|-----|-------|--------------------------|-------|
| | N survivors | | | | | | | | | | |
| Gender, female/male | 49/134 | 14/40 | | 26.7%/73.3% | 25.9%/74.1% | .902 | 0.8% | 235 | 0.12 | -12.7% | 14.4% |
| Age, years | 183 | 54 | 54 | 40.5 \pm 19.9 | 50.8 \pm 22.3 | .001 | -10.4 | 235 | 3.27 | -16.6 | -4.1 |
| GCS | 183 | 54 | 54 | 11.3 \pm 4.2 | 7.5 \pm 4.7 | <.001 | 3.9 | 235 | 5.78 | 2.5 | 5.2 |
| SBP, mmHg | 183 | 51 | 51 | 129 \pm 27 | 124 \pm 38 | .271 | 5.3 | 232 | 1.10 | -4.1 | 14.6 |
| SpO ₂ , % | 180 | 49 | 49 | 95.3 \pm 6.8 | 91.8 \pm 8.7 | .003 | 3.5 | 227 | 3.02 | 1.2 | 5.8 |
| BR, respirations/min | 116 | 31 | 31 | 17.4 \pm 7.7 | 15.0 \pm 8.2 | .130 | 2.4 | 145 | 1.52 | -0.7 | 5.5 |
| HR, beats/min | 175 | 53 | 53 | 92.7 \pm 22.1 | 98.5 \pm 29.5 | .123 | -5.8 | 226 | 1.54 | -13.2 | 1.6 |
| SI | 175 | 50 | 50 | 0.76 \pm 0.28 | 0.85 \pm 0.39 | .055 | -0.10 | 223 | 1.93 | -0.19 | 0.00 |
| Hb, g/dl | 171 | 42 | 42 | 12.3 \pm 2.6 | 10.0 \pm 3.2 | <.001 | 2.3 | 211 | 4.86 | 1.4 | 3.2 |
| PT, % | 161 | 36 | 36 | 78.6 \pm 22.5 | 58.1 \pm 30.1 | <.001 | 20.5 | 195 | 4.63 | 11.7 | 29.3 |
| BASE, mmol/l | 90 | 23 | 23 | -2.9 \pm 5.0 (Median = -3) | -3.9 \pm 7.2 (Median = -5) | .439 | 1.0 | 111 | 0.78 | -1.6 | 3.6 |
| Lactate, mmol/l | 59 | 22 | 22 | 2.9 \pm 2.0 | 5.0 \pm 4.9 | .007 | 2.1 | 79 | 2.78 | 0.6 | 6.4 |
| Hospital stay, days | 183 | 54 | 54 | 15.5 \pm 13.5 | 4.4 \pm 4.6 | <.001 | 11.09 | 235 | 5.91 | 7.4 | 14.8 |
| ISS | 183 | 54 | 54 | 26.3 \pm 7.8 | 40.4 \pm 15.0 | <.001 | -14.2 | 235 | 9.25 | -17.2 | -11.2 |
| T-RTS | 116 | 31 | 31 | 10.5 \pm 1.7 | 8.9 \pm 2.5 | <.001 | 1.61 | 145 | 4.25 | 0.86 | 2.36 |
| PTS | 183 | 54 | 54 | 16.0 \pm 9.7 | 27.2 \pm 14.9 | <.001 | -11.2 | 235 | 6.50 | -14.5 | -7.8 |
| Champion-TRISS | 183 | 54 | 54 | 0.87 \pm 0.19 | 0.50 \pm 0.35 | <.001 | 0.37 | 235 | 10.02 | 0.30 | 0.43 |
| NTDB-TRISS | 183 | 54 | 54 | 0.93 \pm 0.09 | 0.68 \pm 0.27 | <.001 | 0.25 | 235 | 10.60 | 0.20 | 0.30 |
| SAPS II survival | 183 | 54 | 54 | 0.85 \pm 0.19 | 0.45 \pm 0.28 | <.001 | 0.39 | 235 | 11.94 | 0.32 | 0.46 |

Table 3 Logistic regression of prognosis for 30-day-mortality (Model 3 is used for the calculation of the BMTPM)

| Model | Total | | Improvement | | Variance Nagel-kerke R^2 | Variable | B | Wald | P value | Odds ratio | 95% CI Odds ratio | |
|-------|--------------------|---------|--------------------|--------------|-------------------------------|----------|--------|-------|---------|---------------|-------------------|-------|
| | Chi-square (df) | P value | Chi-square (df) | P value | | | | | | | Lower | Upper |
| 1 | 10.01 (1) | .002 | 10.01 (1) | .002 | .063 | Age | .023 | 9.83 | .002 | 1.023 | 1.009 | 1.038 |
| 2 | 79.45 (2) | <.001 | 69.44 (1) | <.001 | .433 | Age | .037 | 15.96 | <.001 | 1.038 | 1.019 | 1.057 |
| 3 | 87.10 (3) | <.001 | 7.65 (1) | .006 | .467 | ISS | .135 | 41.35 | <.001 | 1.144 | 1.098 | 1.193 |
| | | | | | | Age | .044 | 18.74 | <.001 | 1.045 | 1.024 | 1.065 |
| | | | | | | ISS | .116 | 27.12 | <.001 | 1.123 | 1.075 | 1.172 |
| 4a | 92.60 (4) | <.001 | 5.50 | .019 | .491 | GCS | −.129 | 7.42 | .006 | .879 | .802 | .965 |
| | | | | | | Age | 0.043 | 17.38 | <.001 | 1.043 | 1.023 | 1.065 |
| | | | | | | ISS | 0.105 | 22.22 | <.001 | 1.111 | 1.063 | 1.160 |
| 4b | 91.80 (4) | <.001 | 4.70 | .030 | .488 | GCS | −0.136 | 7.94 | .005 | 0.873 | 0.794 | 0.959 |
| | | | | | | Hb | −0.178 | 5.38 | .020 | 0.837 | 0.720 | 0.973 |
| | | | | | | Age | 0.045 | 19.65 | <.001 | 1.046 | 1.025 | 1.067 |
| | | | | | | ISS | 0.111 | 24.83 | <.001 | 1.118 | 1.070 | 1.168 |
| | | | | | | GCS | −0.121 | 6.42 | .011 | 0.886 | 0.806 | 0.973 |
| 5 | 94.00 (5) | <.001 | 1.40 or 2.20 | .237 or .137 | .497 | PT | −1.899 | 4.62 | .032 | 0.150 | 0.026 | 0.847 |
| | | | | | | Age | 0.044 | 18.33 | <.001 | 1.045 | 1.024 | 1.066 |
| | | | | | | ISS | 0.105 | 21.97 | <.001 | 1.111 | 1.063 | 1.161 |
| | | | | | | GCS | −0.130 | 7.1 | <.001 | 0.878 | 0.798 | 0.966 |
| | | | | | | Hb | −0.128 | 2.18 | .140 | 0.880 | 0.743 | 1.043 |
| | | | | | | PT | −1.186 | 1.4 | .236 | 0.305 | 0.043 | 2.173 |

were added to the model combining age, ISS, and GCS, no significant improvement could be demonstrated. With regard to 24-h mortality, age, ISS, and GCS together explained 48.3% of deaths. In addition, SapO₂ contributed significant information, with an overall variance of 53.4% ($P = 0.006$); Hb, with an overall variance of 51.6% ($P = 0.026$); and both parameters together, with an overall variance of 56.7% ($P = 0.002$; detailed models not shown).

On the basis of these findings we drew up the “Basel Multiple Trauma Prediction of Mortality” (BMTPM) for the prognosis of 30-day mortality in our investigation cohort, including age, ISS, and the GCS: $BMTPM = 1 / (1 + (1 / (\text{EXP}(5.686 - 0.044 * \text{age} - 0.116 * \text{ISS} + 0.129 * \text{GCS}))))$. The validation of this score in a similar cohort of 172 multiple trauma patients (in-hospital mortality 22%) previously treated at our institution [49] showed that BMTPM predicted a survival rate of 78.8% with an AUC of 0.805. We tested this score by comparing its discriminative value with that of the two variables showing significance in multivariate analysis as well as the mortality prediction scores under evaluation (Table 4). Figure 1 illustrates the range of the ROC-curves. After we added the hemoglobin and prothrombin variables to the model, the AUC improved from 0.86 to 0.87.

Depending on the score used, predicted mortality ranged from 13% (NTDB-TRISS) to 31% (Bouamra score). Table 4 lists the corresponding Z-scores, which differed significantly above and below the actual observed mortality (BMTPM = 0), i.e., between −5.7 (NTDB-TRISS) and 3.7 (Bouamra score). The actual mortality in our cohort (22.8%) was equally predicted by several TRISS-versions (e.g., Champion-TRISS: mean 0.78, i.e., 22% predicted mortality; MTOS-TRISS: mean 0.77, i.e., 23% predicted mortality) and SAPS II. Of all the models tested, the NTDB-TRISS emerged as the score that combined high precision (AUC 0.84) with the highest benchmark level in the prediction of 30-day mortality (mean 0.87, i.e., 13%). The sensitivity and specificity of all scores under evaluation was lower if elderly or more seriously injured patients were compared to younger or less seriously injured ones (an example of the Champion- and the NTDB-TRISS is given in Table 5).

Discussion

We report three major findings from a search for mortality benchmarking in a prospective cohort of multiply injured patients:

Table 4 Comparison of variables/models predicting 30-day survival (descriptive values and area under the curve, AUC)

| Variable/mortality prediction model | Descriptive values | | | | Area under the curve | | | DEF ^a | |
|-------------------------------------|--------------------|-------|-------|-------|----------------------|------|--------|------------------|--------|
| | Mean | SD | Min | Max | AUC | SE | 95% CI | Z | |
| Age | 42.84 | 20.92 | 7.31 | 91.78 | 0.64 | 0.04 | 0.55 | 0.72 | |
| GCS | 10.44 | 4.60 | 3.00 | 15.00 | 0.73 | 0.04 | 0.65 | 0.81 | |
| ISS | 29.48 | 11.53 | 17.00 | 75.00 | 0.80 | 0.04 | 0.73 | 0.87 | |
| T-RTS | 10.45 | 1.86 | 3.00 | 12.00 | 0.72 | 0.04 | 0.64 | 0.80 | |
| RTS | 6.54 | 1.55 | 1.76 | 7.84 | 0.73 | 0.04 | 0.64 | 0.81 | |
| Boyd-TRISS | 0.79 | 0.28 | 0.00 | 0.99 | 0.83 | 0.03 | 0.77 | 0.89 | −0.89 |
| Champion-TRISS | 0.78 | 0.29 | 0.00 | 0.99 | 0.83 | 0.03 | 0.77 | 0.90 | −0.51 |
| MTOS-TRISS | 0.77 | 0.28 | 0.00 | 0.99 | 0.84 | 0.03 | 0.78 | 0.90 | 0.02 |
| TARN-TRISS | 0.72 | 0.31 | 0.00 | 0.99 | 0.85 | 0.03 | 0.79 | 0.91 | 2.56* |
| NTDB-TRISS | 0.87 | 0.19 | 0.00 | 0.99 | 0.84 | 0.03 | 0.78 | 0.91 | −5.70* |
| Bouamra score | 0.69 | 0.32 | 0.01 | 0.99 | 0.84 | 0.03 | 0.78 | 0.90 | 3.68* |
| SAPS II survival | 0.76 | 0.27 | 0.03 | 0.99 | 0.86 | 0.05 | 0.79 | 0.92 | 0.63 |
| BMPM | 0.77 | 0.25 | 0.01 | 0.99 | 0.86 | 0.03 | 0.80 | 0.92 | 0 |
| 30-day survival (observed) | 0.77 | | | | | | | | |

^a Definitive outcome-based evaluation (DEF): $Z < -1.96 \rightarrow$ survival of study population significantly (*) worse compared with baseline population, $Z > 1.96 \rightarrow$ survival of study population significantly (*) better compared with baseline population

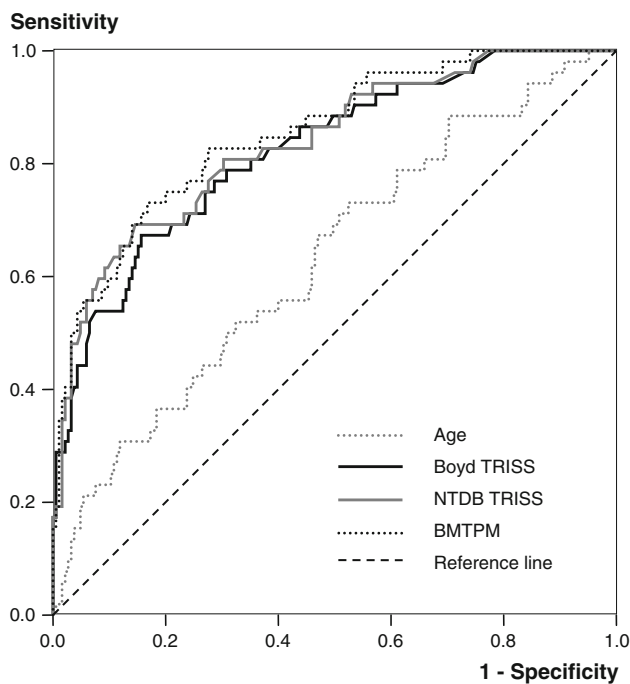


Fig. 1 Comparison of selected models in the prediction of 30-day mortality (ROC curves). ROC Receiver Operating Characteristic curves, TRISS Trauma and Injury Severity Score, in the versions of Boyd (*Boyd-TRISS*) or the National Trauma Data Base (*NTDB-TRISS*), and the Basel Multiple Trauma Prediction of Mortality (*BMPM*)

The actual 30-day mortality rate of 22.8% exactly matched the predictions of the TRISS-versions of Boyd, Champion, and the Major Trauma Outcome Study.

The TRISS using newer National Trauma Data Bank coefficients (TRISS-NTDB) was found to combine the highest requirement profile with precision in the prediction of 30-day mortality.

Our investigation showed that comparison of outcomes with reported values is still almost impossible for a single center not integrated into a trauma registry.

Can we as clinicians be satisfied with our 30-day mortality of 22.8%—a rate that precisely matched the mortality predicted by the early TRISS versions? In 1983 Champion stated that injury severity scales of proven reliability and validity are essential to the accurate prediction of outcome [10]. Quantitative assessment of treatment quality for different multiple trauma cohorts requires well-defined and accepted benchmarking data, e.g., to predict the 30-day mortality rate [5, 8, 50]. Reliable comparison between centers is impossible based on raw or grouped outcome data alone [17, 35]. Without published standardized data, an apparent discrepancy of outcomes may be misleading. Following adjustment for differences in the case mix of individual hospitals, the National Study on the Costs and Outcome of Trauma (NSCOT) showed that the risk of death is significantly lower at a level I trauma center than at a non-trauma center [14]. But, a recent investigation revealed that, in terms of survival, half the American College of Surgeons (ACS)-verified level I trauma centers performed significantly differently from their risk-adjusted expectations [13].

Most workers in the field agree that medical care of multiply injured patients has improved qualitatively in recent decades. The trauma registry of the German Society of Trauma Surgery showed a reduction in mortality from

Table 5 Sensitivity and specificity of Champion-TRISS and NTDP-TRISS in different subsamples (according to Demetriades [47, 59])

| Attribute | <i>N</i> | Survival rate | TRISS | Sensitivity | Specificity | Positive predicted value (PPV) | Negative predicted value (NPV) | Overall classification rate (%) |
|--------------|----------|---------------|----------|-------------|-------------|--------------------------------|--------------------------------|---------------------------------|
| Total sample | 237 | .78 | Champion | 0.93 | 0.54 | 0.69 | 0.87 | 0.84 |
| | | | NTDP | 0.98 | 0.19 | 0.77 | 0.80 | 0.80 |
| ISS ≤ 25 | 107 | .91 | Champion | 0.99 | 0.10 | 0.50 | 0.91 | 0.91 |
| | | | NTDP | 1.00 | 0.00 | n.a. | 0.91 | 0.91 |
| ISS > 25 | 130 | .68 | Champion | 0.86 | 0.64 | 0.70 | 0.82 | 0.78 |
| | | | NTDP | 0.97 | 0.23 | 0.77 | 0.71 | 0.72 |
| Head AIS ≤ 2 | 98 | .90 | Champion | 0.99 | 0.17 | 0.67 | 0.89 | 0.89 |
| | | | NTDP | 1.00 | 0.00 | n.a. | 0.88 | 0.88 |
| Head AIS > 2 | 139 | .70 | Champion | 0.88 | 0.64 | 0.69 | 0.85 | 0.81 |
| | | | NTDP | 0.97 | 0.24 | 0.77 | 0.75 | 0.75 |
| GCS > 3 | 201 | .84 | Champion | 0.99 | 0.29 | 0.83 | 0.87 | 0.87 |
| | | | NTDP | 0.99 | 0.06 | 0.50 | 0.84 | 0.83 |
| GCS = 3 | 36 | .44 | Champion | 0.31 | 0.95 | 0.63 | 0.83 | 0.67 |
| | | | NTDP | 0.94 | 0.40 | 0.89 | 0.56 | 0.64 |
| Age < 55 | 168 | .82 | Champion | 0.93 | 0.58 | 0.64 | 0.91 | 0.86 |
| | | | NTDP | 0.99 | 0.10 | 0.75 | 0.83 | 0.83 |
| Age ≥ 55 | 69 | .68 | Champion | 0.93 | 0.48 | 0.79 | 0.78 | 0.78 |
| | | | NTDP | 0.96 | 0.30 | 0.78 | 0.73 | 0.74 |

n.a. not analyzed (as specificity = 0)

1993 to 2005 from 22.8% to 18.7% [51]. In contrast, other European hospitals have reported unchanged or even increased mortality rates over the last 10–15 years, although ISS and trauma patterns have remained comparable [52]. Not only the quality of treatment may change over time: In recent years centers in Austria and the Netherlands have reported a decreasing mean ISS, an increased mean patient age, and a higher percentage of severe head injuries among the multiply injured [16, 53]. Given these discrepant and not uniformly adjusted data on mortality [18], individual centers can hardly find reliable data with which to compare their own performance and evaluate their treatment of multiple trauma patients. Participation in national multicenter databases may offer regional solutions, with the potential advantage of eliminating geographic variations in patient characteristics, the epidemiology of injury, or the health care provided. Nevertheless, from an international perspective, registers using different inclusion criteria and their own optimized assessment scores only fail to permit international comparison [54–56].

Our study population is representative of a typical European multiple trauma cohort in the upper range of trauma severity [9, 16, 57], with a mean age of 43 years, an ISS of 30, and a large majority of blunt injuries following traffic accidents. As for studies based on the German Trauma Registry [51] the focus of our investigation was on the primary treatment of multiple trauma only and excluded patients with mono- or minor trauma as well as those

admitted secondarily in order to avoid major confounders. With the objective of appraising the raw 30-day mortality rate in our cohort of multiple trauma patients adequately, we tested several prediction models published in the literature, e.g., different variations of the TRISS. We came to realize that several authors had not stated which version of the TRISS they actually used [44, 46, 58, 59]. The TRISS and the coefficients it includes were originally developed in the 1980 s by Boyd et al. based on the US-American Major Trauma Outcome Study (MTOS) database for motor vehicle accidents [11]. Since then, the need to update or adjust these coefficients has been discussed [60, 61] and various modifications of the original Boyd-TRISS have been published in an effort to predict mortality more precisely [23, 24, 44, 60]. Despite some important criticisms and the subsequent development of several other prediction models [51, 57, 58, 62], the TRISS still figures as the internationally most frequently applied instrument for the assessment and adjustment of injury severity facilitating comparison of survival rates for (multiple) trauma patients [24, 44, 55, 63].

Our actual rate of non-survival (22.8%) was equally well predicted by the Champion [23]- and MTOS [23]-TRISS (AUC 0.83 and 0.84, respectively). From a statistical point of view and given the minimal differences in predicted survival and precision (AUC values) of these models, as well as the observed deviations in comparison to the similar Boyd-TRISS or SAPS II, we cannot advocate any

single outperforming score. In 1997 the Cologne Validation Study showed that both the original Boyd and the subsequent Champion versions of the TRISS were highly accurate for predicted mortality, even though trauma patients and care are different in the US and Germany [9]. Other European reports on trauma populations not confined to multiple trauma support this finding for the Boyd-TRISS [64]; others observed a better than expected survival rate, regardless of whether newer MTOS- or TARN coefficients were used [44]. The fact that the TRISS can identify major differences in levels of multiple trauma care is clear from observed mortality rates in developing countries that are two or three times higher than the TRISS mortality prediction [46, 65]. In contrast, an unrealistic >50% rate of high-performance hospitals has been reported regarding TRISS and ASCOT-predicted mortality data using older MOTS- formulae [12]. An updated TRISS with new NTDB-derived coefficients yielded an improved mortality prediction compared to the older MTOS [66]-TRISS [55].

Against the background of this debate, we tested the correlation of routine emergency parameters with mortality, in a first step, and the predictive ability of the resultant combined scores, in a second step. As expected, combined scores outperformed the single emergency parameters in the prediction of mortality. The treatment-restricted variables age, ISS, and GCS were found to be the main independent prognostic parameters in our multivariate analysis. The prominent importance of each of these three variables, all being part of the TRISS, has been described before [67–69]. Historically, the inclusion of additional parameters into any prediction model attracted heightened clinical interest if these were associated with a potentially therapeutic option, such as the correction of base excess, or other specific optimization of treatment [21, 54]. However, other studies found that the existence of a cause-and-effect relationship could not be proven even when several risk factors were clearly associated with poorer outcome [2, 68]. In contrast to other investigators [4, 35, 70], we found that, in addition to these three variables, only the initial hemoglobin value added some small prognostic information on both 24-h and 30-day mortality in multivariate analysis. Initial prothrombin time appeared to have a comparable impact on 30-day mortality only. Our observation that this effect was no longer statistically significant if the hemoglobin and prothrombin times were both added to the model provides indirect evidence for the strong association of these two measurements. Of course, this finding makes sense from a clinical point of view and supports published reports on blood loss and traumatic coagulopathy [19, 54, 71]. Unfortunately, the fact that we could not adequately record transfusion volume requirements or pre-trauma anticoagulant status meant that we had to forego interpretation of these two values (i.e., 10.1% for

hemoglobin and 16.9% for prothrombin). In addition, the limited statistical power of our cohort in comparison to larger databases may explain why we could not clearly identify any additional variables of influence.

In recent years numerous prediction models have been described in the literature, yielding conflicting data in terms of constitution and precision. A prognostic model including base excess and prothrombin time as significantly independent factors in addition to age, ISS, and GCS reached an AUC of 0.90 in the German Trauma Registry [72, 73]. Lackner et al. confirmed this finding for another center also participating in that registry [70], but with a lower AUC (0.82). Astonishingly, even though we did not include the two variables base excess and prothrombin time in our prognostic BMTPM-model we observed a higher precision for the prediction of mortality (AUC 0.86). A disadvantage of adding various variables to any existing prediction model is that it becomes more complicated and, in effect, more prone to dropout in terms of missing values. The two parameters base excess and prothrombin time are particularly susceptible to a high missing rate in the emergency situation [19, 74], which may then negate the enhanced precision achieved by adding the variables. Furthermore, adding hemoglobin and prothrombin added only marginal precision to the BMTPM-model (+0.01 AUC). The relatively simple BMTPM, with its advantage of avoiding missing breathing values, was found to have the highest prognostic value in comparison to all other variables tested in our study, and it was equal to or even better than standard scores like the TRISS or SAPS II, which include more parameters but are challenged by the problem of missing values. Of course, a high prognostic value will be expected because the BMTPM was developed in this specific cohort. Our finding of a lower AUC when testing the model in a retrospective cohort supports this expectation. The equally precise SAPS II score has the disadvantage that it needs many more subvariables for determination.

A few authors report even higher precision for their prediction models, e.g., an AUC of 0.91 for the SAPS II, 0.95 for the Bouamra score [57] or an AUC of 0.97 for a version of the TRISS that is not clearly specified [58]. In comparison with our investigation, the observed differences might generally be explained by the lack of adjustment for different study cohorts. For example, Demetriades et al. demonstrated a higher misclassification rate when applying TRISS methodology to elderly patients or those suffering more severe trauma [47, 59], and as a result, they concluded that this approach should be discontinued. Even though we also observed reduced sensitivity and specificity in mortality prediction for all the scores under investigation in the relevant patient subgroups, their final conclusion has to be critically reviewed. From a statistical perspective, the discriminatory power of any scoring system to predict

mortality will be increased if more extreme outcome cases are included in the analysis. In large general trauma databases with predominantly mild trauma patients, there is always an overwhelming probability (>95%) of survival, in contrast to an almost negligible probability of non-survival. In this light and given the complexity of most prediction scores, it is debatable whether the application of a specificity model to favor selected subgroup analysis is plausible in clinical routine.

Overall, in our collective, the quality of the prediction models under investigation differed only marginally between AUC 0.82 and 0.86. What did differ was the predicted mortality rate of the corresponding scores, which ranged from 13% (NTDB-TRISS) to 31% (Bouamra score). The Bouamra score with its complex methodology includes gender as well as age, ISS, and GCS, but even so, yielding an almost identical AUC in comparison to all other models in our cohort clearly overestimated mortality. For the purpose of quality assurance the equally precise NTDB-TRISS with its lower predicted mortality provided the highest target benchmark. A more simplistic approach would be just to aspire to obtaining a higher survival rate than that predicted by any of the prediction models, but this carries the disadvantage of not knowing the extent of success. As discussed earlier, the evidence that the NTDB-TRISS provides a more adequate assessment of the level of performance aspired to today favors our approach [12]. Consequently, our mortality rate of 9% higher than predicted with the newer NTDB-TRISS supports the conclusion that our treatment outcomes can be improved.

Apart from this challenging clinical implication, our investigation illustrates the need for a well-accepted international benchmarking process in the assessment of mortality for the multiply injured patient. Understandably, clinicians are concerned about whether any such method actually reflects quality of care, especially if trauma centers are to be ranked by mortality rate [13, 75]. In our opinion, such quantitative statistical outcome analysis will only be a first step in the quality assessment of multiple trauma therapy at individual centers. Undoubtedly, detailed qualitative analysis of critical clinical processes as well as case report forums, such as interdisciplinary morbidity and mortality conferences, remain indispensable for achieving any further improvement of treatment [47, 76].

This study has several limitations:

We restricted our investigation primarily to 30 day-mortality. This prospective consecutive series represents a single-center experience with the disadvantage of a limited case number, and it cannot compete with large national databases. Our perspective was different because we wished to assess quality of performance in an institution that is not part of a larger registry by comparing actual

mortality to reported values. Because our study cohort included multiply injured patients only, statistical comparison with general trauma registries such as MTOS, NTDB, or TARN was not possible due to the missing match of study populations. Independent of this fact, any further matching procedures, including *M*-, *Z*- and *W*-statistics, are highly sophisticated and hardly realistic for single centers that are not integrated into large national databases [45, 48, 63]. In contrast to other researchers, we did not exclude either major head trauma [30], patients who did not survive the first 24 h [77] or patients for whom some values were missing [55]. Our analysis was restricted to the parameters that are routinely obtained in the treatment of multiply injured patients, and therefore we cannot comment on other variables frequently discussed in this context, such as cytokines or inflammation markers.

Unlike other studies [78, 79], we did not find any impact of gender on outcome. As is typical for a mid-European hospital, we experienced a very low rate of penetrating trauma, which limits comparison with centers reporting on a much higher incidence of violence-related trauma [9]. Similarly, we cannot comment on the potential impact of race on outcome as reported, for example, in studies from the United States. [80]. We did not investigate the effects of variables such as co-morbidity, medication, or transfusion requirements, all of which are under debate as major contributors to outcome but which are difficult to obtain and to score in a reliable manner in the emergency setting of multiple trauma [54, 71, 81]. Because our study focused on the prediction models discussed here, we cannot comment on the effect of other scores or stratification methods [20, 46, 55, 82].

Another well-known problem of any clinical study analyzing numerous parameters is the handling of missing values, especially in the emergency treatment period, e.g., lactate, breath rate or the resultant scores. Some of the reported differences in the accuracy of prediction models may, in reality, be due to variable completeness of data acquisition for the different study cohorts. The trauma registry of the German Society of Trauma Surgeons reports a 39% rate of cases with missing base excess data [74], and a large Canadian trauma database reports a missing rate of almost 40% of GCS data and 5.6% of respiration rate values [40]. The latter accounted for 22% of deaths [32]. The exclusion of patients with missing values from further analysis [55] may not only diminish the number of cases analyzed but also, and even more importantly, may create an undesirable bias, possibly related to the specific composition of any particular cohort [83]. We observed an important number of missing values for base excess, lactate, and respiration rate, with the consequence that it was impossible to calculate the RTS and TRISS for those patients. The subgroup of patients with missing values did

not show any significant differences from the rest of the cohort with regard to the variables tested. To compensate for this problem, we decided to uniformly impute neutral values to substitute for missing cases [44]. The imputation method has been proven to yield hospital quality measures that are almost identical to those based on the true data [32, 83, 84]. Other groups tried to solve the problem of missing data by choosing a mortality prediction model in which variables with a higher rate of missing values, such as the RTS, were substituted by another variable with low or no missing data, for example, the GCS [42, 85]. In the light of the finding that Bouamra scoring did not perform as well as other TRISS versions in our study cohort, we cannot advocate this preference.

Other inconsistencies in the way raw data are handled by different investigators in emergency medicine concern the time at which ISS diagnosis is made, i.e., on patient admission or at the end of the hospital stay. We decided in favor of the latter because of the well-known tendency toward early underscoring of trauma in emergency situations. In addition, differences may arise depending on whether field values or emergency department measurements are used for the first evaluation and subsequent scoring of patients [85].

Overall, Clark has reasoned that, because there are so many possible limitations, the comparison of institutional trauma survival to a standard value will continue to be a challenge [86].

In summary, with the aim of assessing our institution's quality of performance in the treatment of multiply injured patients, we tried to compare it with the mortality rates reported in the literature. Lacking an international standard, we needed to find out which of the investigated prediction models would emerge as the most useful for clinical routine. In our investigation, despite all constraints, the NTDB-TRISS appeared to provide the best score, simultaneously combining high statistical precision with the highest therapeutic aims in terms of patient survival after multiple trauma. For today's single center without the chance of participating in a large, high-quality trauma registry, we recommend the use of the NTDB-TRISS as an important first step toward establishing a critical quality assessment of the treatment of multiple trauma patients. Our investigation underlines the clinical need for a benchmarking procedure, accepted worldwide, that would be an integral part of standardized quality management for all multiply injured patients.

Acknowledgments This study was financially supported by the Swiss National Science Foundation (NCCR-CO-ME), the Scientific Foundation of the Swiss Accident Insurance Fund (SUVA), the Voluntary Academic Society (FAG) Basel, the Swiss group of Computer Assisted Radiology and Surgery (CARCAS) Basel, the Unit of Interventional Radiology at the University Hospital Basel, and

the Swiss National Information Center for Accident Prevention (BFU). The authors thank J. Buchanan for editorial assistance with the manuscript, J. Reinkensmeier for help creating the computer database, and all the patients and collaborators involved in the study at the University Hospital Basel.

References

1. Uranues S, Lamont E (2008) Acute care surgery: the European model. *World J Surg* 32:1605–1612
2. Sauer A, Moore F, Moore E et al (1995) Epidemiology of trauma deaths: a reassessment. *J Trauma* 38:185–193
3. Neary W, Prytherch D, Foy C et al (2007) Comparison of different methods of risk stratification in urgent and emergency surgery. *Br J Surg* 94:1300–1305
4. Nathens A, Xiong W, Shafi S (2008) Ranking of trauma center performance: the bare essentials. *J Trauma* 65:628–635
5. Budd H, Almond L, Oakley P et al (2008) A benchmarking study of two trauma centers highlighting limitations when standardising mortality for comorbidity. *World J Emerg Surg* 3:2
6. Neugebauer E, Lefering R (2002) Severity scores in surgery: what for and who needs them? An introduction: definition, aims, classification and evaluation. *Langenbecks Arch Surg* 387: 55–58
7. Willcox N, Oakley P (2002) Survival with an arterial pH of 6.57 following major trauma with exsanguinating haemorrhage associated with traumatic amputation. *Resuscitation* 53:217–221
8. Skaga N, Eken T, Steen P (2006) Assessing quality of care in a trauma referral center: benchmarking performance by TRISS-based statistics or by analysis of stratified ISS data? *J Trauma* 60:538–547
9. Bouillon B, Lefering R, Vorweg M et al (1997) Trauma score systems. *J Trauma* 42:652–658
10. Champion H, Sacco W, Hunt T (1983) Trauma severity scoring to predict mortality. *World J Surg* 7:4–11
11. Boyd C, Tolson M, Copes W (1987) Evaluating trauma care: the TRISS method. *J Trauma* 27:370–378
12. Glance L, Osler T, Dick A (2005) Evaluation trauma center quality: does the choice of the severity-adjustment model make a difference? *J Trauma* 58:1265–1271
13. Shafi S, Nathens A, Parks J et al (2008) Trauma quality improvement using risk-adjusted outcomes. *J Trauma* 64:599–606
14. MacKenzie E, Rivara F, Jurkovich G et al (2006) A national evaluation of the effect of trauma-center care on mortality. *N Engl J Med* 354:366–378
15. Scarborough K, Slone D, Uribe P et al (2008) Reduced mortality at a community hospital trauma center: the impact of changing trauma level designation from II to I. *Arch Surg* 143:22–28
16. Nijboer J, van der Sluis C, van der Naalt J et al (2007) Two cohorts of severely injured trauma patients, nearly two decades apart: unchanged mortality but improved quality of life despite higher age. *J Trauma* 63:670–675
17. Bauzà G, LaMorte W, Burke P et al (2008) High mortality in elderly drivers is associated with distinct injury patterns: analysis of 187, 869 injured drivers. *J Trauma* 64:304–310
18. Christensen M, Nielsen T, Ridley S et al (2008) Outcomes and costs of penetrating trauma injury in England and Wales. *Injury* 39:1013–1025
19. Maegele M, Lefering R, Yucel N et al (2007) Early coagulopathy in multiple injury: an analysis from the German Trauma Registry on 8724 patients. *Injury* 38:298–304
20. Cheng C, Graham C, Gabbe B et al (2008) Trauma care systems—a comparison of trauma care in Victoria, Australia, and Hong Kong, China. *Ann Surg* 247:335–342

21. Macleod J, Lynn M, McKenney M et al (2004) Predictors of mortality in trauma patients. *Am Surg* 70:805–810
22. Clark D, Lucas F, Ryan L (2007) Predicting hospital mortality, length of stay, and transfer to long-term care for injured patients. *J Trauma* 62:592–600
23. Champion H, Copes W, Sacco W et al (1990) A new characterization of injury severity. *J Trauma* 30:539–546
24. Millham F, LaMorte W (2004) Factors associated with mortality in trauma: re-evaluation of the TRISS method using the National Trauma Data Bank. *J Trauma* 56:1090–1096
25. Vance S (2007) The FAST scan: are we improving care of the trauma patient? *Ann Emerg Med* 49:364–366
26. Committee on Trauma, American College of Surgeons (1997) Advanced trauma life support program for doctors. American College of Surgeons, Chicago
27. Gennarelli T, for the American Association for Automotive Medicine (1990) The abbreviated injury scale. American Association for Automotive Medicine, Des Plaines, IL
28. Ruchholtz S (2000) The Trauma Registry of the German Society of Trauma Surgery as a basis for interclinical quality management. A multicenter study of the German Society of Trauma Surgery. *Unfallchirurg* 103:30–37
29. Teasdale G, Jennett B (1974) Assessment of coma and impaired consciousness. *Lancet* 2:81–83
30. Eastridge B, Salinas J, McManus J et al (2007) Hypotension begins at 110 mmHg: redefining “hypotension” with data. *J Trauma* 63:291–299
31. Belzberg H, Wo C, Demetriades D et al (2007) Effects of age and obesity on hemodynamics, tissue oxygenation, and outcome after trauma. *J Trauma* 62:1192–1200
32. Moore L, Lavoie A, LeSage N et al (2006) Statistical validation of the revised trauma score. *J Trauma* 60:305–311
33. Chien L-C, Lu K, Wo C et al (2007) Hemodynamic patterns preceding circulatory deterioration and death after trauma. *J Trauma* 62:932
34. Grimme K, Pape H, Probst C et al (2005) Calculation of different triage scores based on the German Trauma Registry. *Eur J Trauma* 31:480–487
35. Lichtveld R, Panhuizen I, Smit R et al (2007) Predictors of death in trauma patients who are alive on arrival at hospital. *Eur J Trauma Emerg Surg* 33:46–51
36. Mikulaschek A, Henry S, Donovan R et al (1996) Serum lactate is not predicted by anion gap or base excess after trauma resuscitation. *J Trauma* 40:218–224
37. Kaplan L, Frangos S (2005) Clinical review: acid-based abnormalities in the intensive care unit. *Crit Care* 9:198–203
38. Baker S, O'Neill B, Haddon W et al (1974) The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma* 14:187–196
39. Falcone RE, Santanello SA, Schulz MA et al (1993) Correlation of metabolic acidosis with outcome following injury and its value as a scoring tool. *World J Surg* 17:575–579
40. Moore L, Lavoie A, Abdous B et al (2006) Unification of the revised trauma score. *J Trauma* 61:718–722
41. Oestern H, Tscherne H, Sturm J et al (1985) Classification of the severity of injury. *Unfallchirurg* 88:465–472
42. Bouamra O, Wrotchford A, Hollis S et al (2006) Outcome prediction in trauma. *Injury* 37:1092–1097
43. Le Gall J, Lemeshow S, Saulnier F (1993) A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. *JAMA* 270:2957–2963
44. Hansen K, Engesaeter L, Eide G et al (2005) Trauma outcome in a Norwegian regional hospital. *Injury Extra* 36:361–367
45. Hollis S, Yates D, Woodford M et al (1995) Standardized comparison of performance indicators in trauma: a new approach to case-mix variation. *J Trauma* 38:763–766
46. Murlidhar V, Roy N (2004) Measuring trauma outcome in India: an analysis based on TRISS methodology in a Mumbai university hospital. *Injury* 35:386–390
47. Demetriades D, Chan L, Velmanos G et al (2001) TRISS methodology: an inappropriate tool for comparing outcomes between trauma centers. *J Am Coll Surg* 193:250–254
48. Younge P, Coats T, Gurney D et al (1997) Interpretation of the Ws statistic: application to an integrated trauma system. *J Trauma* 43:511–515
49. Hirschmann M, Uike K, Kaufmann M et al (2007) Quality management of interdisciplinary treatment of polytrauma. Possibilities and limits of retrospective routine data collection. *Der Anaesthesist* 56:673–678
50. Skaga N, Eken T, Jones J et al (2008) Different definitions of patient outcome: consequences for performance analysis in trauma. *Injury* 39:612–622
51. Ruchholtz S, Lefering R, Paffrath T et al (2008) Reduction in mortality of severely injured patients in Germany. *Dtsch Arztebl Int* 105:225–231
52. Aldrian S, Nau T, Koenig F et al (2006) Management of polytrauma patients during 10 years. *Osteo Trauma Care* 14:93–97
53. Aldrian S, Koenig F, Weninger P et al (2007) Characteristics of polytrauma patients between 1992 and 2002: what is changing? *Injury* 38:1059–1064
54. Dutton R, Lefering R, Lynn M (2006) Database predictors of transfusion and mortality. *J Trauma* 60:S70–S77
55. Kilgo P, Meredith J, Osler T (2006) Incorporating recent advances to make the TRISS approach universally available. *J Trauma* 60:1002–1009
56. The EuroTARN Writing Committee on behalf of the EuroTARN Group (2007) A comparison of European trauma registries: the first report from the EuroTARN Group. *Resuscitation* 75:286–297
57. Bouamra O, Wrotchford A, Hollis S et al (2006) A new approach to outcome prediction in trauma: a comparison with the TRISS model. *J Trauma* 38:185–193
58. Guzzo J, Bochicchio G, Napolitano L et al (2005) Prediction of outcomes in trauma: anatomic or physiologic parameters? *J Am Coll Surg* 201:891–897
59. Demetriades D, Chan L, Velmahos G et al (1998) TRISS methodology in trauma: the need for alternatives. *Br J Surg* 85:379–384
60. Hannan E, Mendeloff J, Farrell L et al (1995) Validation of TRISS and ASCOT using a non-MTOS trauma registry. *J Trauma* 38:83–88
61. Wisner D (2006) Editorial comment. *J Trauma* 61:722
62. Rutledge R, Osler T, Emery S et al (1998) The end of the Injury Severity Score (ISS) and the Trauma and Injury Severity Score (TRISS): ICISS, in international classification of diseases, ninth revision-based prediction tool, outperforms both ISS and TRISS as predictors of trauma patient survival, hospital charges, and hospital length of stay. *J Trauma* 44:41–49
63. Kirkham J, Bouamra O (2008) The use of statistical process control for monitoring institutional performance in trauma care. *J Trauma* 65:1494–1501
64. David J, Gelas-Dore B, Inaba K et al (2007) Are patients with self-inflicted injuries more likely to die? *J Trauma* 62:1495–1500
65. Chardoli M, Rahimi-Movaghar V (2006) Analysis of trauma outcome at a university hospital in Zahedan, Iran using the TRISS method. *East Afr Med J* 83:440–442
66. Champion H, Copes W, Sacco W et al (1990) The major trauma outcome study: establishing national norms for trauma care. *J Trauma* 30:1356–1365
67. McKinley B, Marvin R, Cocanour C et al (2000) Blunt trauma resuscitation: the old can respond. *Arch Surg* 135:688–695
68. Velmahos G, Demetriades D, Shoemaker W et al (2000) Endpoints of resuscitation of critically injured patients: normal or supranormal? *Ann Surg* 232:409–418

69. Linn S, Levi L, Grunau P et al (2007) Effect measure modification and confounding of severe head injury mortality by age and multiple organ injury severity. *Ann Epidemiol* 17:142–147
70. Lackner C, Burghofer K, Stolpe E et al (2007) Prognostic value of routine parameters and laboratory parameters after major trauma. *Der Unfallchirurg* 4:307–319 [in German]
71. Eastridge B, Malone D, Holcomb J (2006) Early predictors of transfusion and mortality after injury: a review of the data-based literature. *J Trauma* 60:S20–S25
72. Rixen D, Raum M, Bouillon B et al (2001) Predicting the outcome in severe injuries: An analysis of 2069 patients from the trauma register of the German Society of Traumatology (DGU). *Unfallchirurg* 104:230–239 (in German)
73. Matthes G, Seifert J, Bogatzki S et al (2005) Age and prognosis of polytrauma patients. Local tailoring of the DGU prognosis model. *Unfallchirurg* 108:288–292 (in German)
74. Rixen D, Raum M, Bouillon B et al (2002) The base excess as a prognostic indicator in polytrauma patients. *Anesthesiol Intensivmed Notfallmed Schmerzther* 37:347–349 (in German)
75. Burney R (2008) Editorial comment. *J Trauma* 65:634–635
76. Teixeira P, Inaba K, Hadjizacharia P et al (2007) Preventable or potentially preventable mortality at a mature trauma center. *J Trauma* 63:1338–1347
77. Martin M, FitzSullivan E, Salim A et al (2006) Discordance between lactate and base deficit in the surgical intensive care unit: which one do you trust? *Am J Surg* 191:625–630
78. Holbrook T, Hoyt D, Anderson J (2001) The importance of gender on outcome after major trauma: functional and psychological outcomes in women versus men. *J Trauma* 50:270–273
79. Harbrecht BG, Peitzman AB, Rivera L et al (2001) Contribution of age and gender to outcome of blunt splenic injury in adults: multicenter study of the Eastern Association for the Surgery of Trauma. *J Trauma* 51:887–895
80. Haider A, Chang D, Efron D et al (2008) Race and insurance status as risk factors for trauma mortality. *Arch Surg* 143:945–949
81. Sacco W, Copes W, Bain LJ et al (1993) Effect of preinjury illness on trauma patient survival outcome. *J Trauma* 35:538–542
82. Osler T, Gance L, Buzas J et al (2008) A trauma mortality prediction model based on the anatomic injury scale. *Ann Surg* 247:1041–1048
83. Gance L, Osler T, Mukamel D et al (2009) Impact of statistical approaches for handling missing data on trauma center quality. *Ann Surg* 249:143–148
84. Newgard C (2006) The validity of using multiple imputation for missing out-of-hospital data in a state trauma registry. *Acad Emerg Med* 13:314–324
85. Arbabi S, Jurkovich G, Wahl W et al (2004) A comparison of prehospital and hospital data in trauma patients. *J Trauma* 56:1029–1032
86. Clark D (1999) Comparing institutional trauma survival to a standard: current limitations and suggested alternatives. *J Trauma* 47:S92–S98