The feasibility, appropriateness, and applicability of trauma scoring systems in low and middle-income countries: a systematic review

Isabelle Feldhaus,1 Melissa Carvalho,2 Ghazel Waiz,3 Joel Igu4 Zachary Matthay,3 Rochelle Dicker,1 Catherine Juillard2

ABSTRACT

Background About 5.8 million people die each year as a result of injuries, and nearly 90% of these deaths occur in low and middle-income countries (LMIC). Trauma scoring is a cornerstone of trauma quality improvement (QI) efforts, and is key to organizing and evaluating trauma services. The objective of this review was to assess the appropriateness, feasibility, and QI applicability of traditional trauma scoring systems in LMIC settings.

Materials and methods This systematic review searched PubMed, Scopus, CINAHL, and trauma-focused journals for articles describing the use of a standardized trauma scoring system to characterize holistic health status. Studies conducted in high-income countries (HIC) or describing scores for isolated anatomic locations were excluded. Data reporting a score’s capacity to discriminate mortality, feasibility of implementation, or use for QI were extracted and synthesized.

Results Of the 896 articles screened, 336 were included. Over half of studies (56%) reported Glasgow Coma Scale, followed by Injury Severity Score (ISS; 51%), Abbreviated Injury Scale (AIS; 24%), Revised Trauma Score (RTS; 19%), Trauma and Injury Severity Score (TRISS; 14%), and Kampala Trauma Score (7%). While ISS was overwhelmingly predictive of mortality, 12 articles reported limited feasibility of ISS and/or AIS. RTS consistently underestimated injury severity. Over a third of articles (37%) reporting TRISS assessment observed mortality that was greater than that predicted by TRISS. Several articles cited limited human resources as the key challenge to feasibility.

Conclusions The findings of this review reveal that implementing systems designed for HICs may not be relevant to the burden and resources available in LMICs. Adaptations or alternative scoring systems may be more effective.

PROSPERO registration number CRD42017064600.

INTRODUCTION

About 5.8 million people die each year as a result of injuries, and nearly 90% of these deaths occur in low and middle-income countries (LMIC). Improvements in the organization of trauma care have resulted in significant reductions in mortality in high-income countries (HIC), such as the USA and Canada. Likewise, there have been increasing efforts to improve trauma care management in LMICs.

Trauma scoring is a cornerstone of trauma care improvement interventions. Standardized trauma scoring systems allow for appropriate triage and classification of trauma patients as well as prediction of patient outcomes and risk adjustment when evaluating patient outcomes and hospital performance. They enable programs to highlight opportunities for trauma care improvement, including among such as identification of observed mortality among patients with low injury severity that may be higher than predicted, as well as to compare trauma outcomes against established norms. They can thus provide an objective definition for preventable deaths and indicate probability of survival, acting as the foundation for mortality review conferences, audits, filters, and other quality improvement (QI) initiatives.

Several trauma scoring systems have been developed and validated in the past 50 years, including the Injury Severity Score (ISS), Abbreviated Injury Scale (AIS), Revised Trauma Score (RTS), and Trauma and Injury Severity Score (TRISS), among others (figure 1). Each type of score was developed with distinct purposes across the continuum of care. The majority of trauma scoring systems were developed for resource-rich settings, however, and often require regular access to advanced diagnostics. Few have been validated for use in LMICs. In fact, studies consistently demonstrate that many scoring systems underpredict mortality in these settings. As a result, physiological-based systems, such as the Kampala Trauma Score (KTS), have been developed and shown to be valid triage tools to predict patient outcomes and inform decision-making.

Despite evidence that certain trauma scoring systems may not be appropriate for resource-limited settings, those commonly found in HICs may also be used in LMICs because they remain the most well established and accepted for trauma management. Yet, the extent to which each scoring system is used in LMICs around the world remains unknown. Understanding which and how trauma scoring systems are used is the first step in determining which score(s) may be optimal in these settings. The optimal trauma score for any setting will strike the appropriate balance between accuracy in quantifying injury severity and feasibility in implementation. The objective of this review was to identify scores that have been used in LMICs to classify injury severity and assess their...
appropriate assessment, feasibility, and application in QI efforts in these settings.

MATERIALS AND METHODS

Search strategy

Electronic databases PubMed, Scopus, and CINAHL as well as selected academic journals were searched for relevant articles published since 1990. The final database search was conducted on November 13, 2017. Search queries used MeSH (Medical Subject Headings) and keywords, including ‘injury’, ‘trauma’, ‘severity’, ‘score’, ‘low and middle income countries’ and score-specific terms (online supplementary appendix 1). Countries were classified as LMICs according to the World Bank country income classifications.16 Articles describing the use of a standardized trauma or injury scoring system to characterize holistic health status were eligible for inclusion. Studies were limited to those published in English or French. Studies conducted in HICs were not included.

Articles describing the use of scoring systems for isolated organ systems and/or anatomic locations were excluded. Articles describing the use of Glasgow Outcome Scale (GOS) without reference to the use of Glasgow Coma Scale (GCS) were not included in this review, since GOS is primarily used to classify patients with brain injuries. GCS is used to evaluate overall consciousness as a component of trauma scoring systems and was, therefore, included in this review. Emotional trauma and psychiatric-related scoring systems were excluded.

Titles and abstracts of search results were screened for eligibility and the selected full-text articles were obtained. Articles that stated that no trauma scoring system was available were noted as a proxy for feasibility, though not included in data analyses. The references of included articles were screened to identify additional articles not initially captured in initial searches to ensure robustness of the study. Retrieved records were imported into Microsoft Excel 2016 and Mendeley V1.17.10.

Data extraction and analysis

Two authors independently reviewed each full-text article and extracted the following information: publication year, country setting, study objective, study design, sample size, and trauma score(s) reported. Studies were categorized into: (1) those reporting or using a trauma score without further evaluation of the system, and (2) those evaluating the trauma scoring system. Articles in the second category underwent a quality (bias) assessment using a standardized critical appraisal checklist (online supplementary appendix 2) adapted from Fowkes and Fulton and Liberman et al.17 18 Discrepancies between reviewers were resolved by discussion and consensus. In the event that no consensus was reached, a third party acted as an arbiter.

Descriptive statistics of study characteristics were generated by country, world region, and type of scoring system. For articles reporting evaluations of scoring systems, data on a trauma score’s capacity to discriminate mortality, feasibility of implementation, or use in QI efforts were extracted. Meta-analysis was deemed inappropriate due to the wide diversity of primary outcomes and study designs among the articles included for review.

RESULTS

Of the 896 articles screened, 336 reporting trauma scores in LMICs were included in this review (figure 2). The reporting of trauma scores in published studies generally increased over time, except in recent years (figure 3). Among the 50 countries represented, a majority of articles focused on South Africa (52 studies), followed by Nigeria (40 studies), India (28 studies), Iran (28 studies), and Malaysia (26 studies). Studies most often focused on upper middle-income countries (163 studies; 49%) and lower middle-income countries (135 studies; 40%), whereas low-income settings composed only 16% (53 studies) of reviewed articles (figure 4).

Over half of studies (187 studies; 56%) reported GCS among their assessments of trauma severity, followed by ISS (173 studies; 51%), AIS (81 studies; 24%), RTS (65 studies; 19%), and TRISS (47 studies; 14%). KTS (24 studies; 7%) was reported...
least often. Other miscellaneous scoring systems, such as indirect stratification based on urgency, victim perception, or other proxy measures, were reported in 11% (37 studies) of articles. Over 33% (112 studies) of articles reported quantitative results on at least one of these features, most commonly reporting differences in score means between population groups, ORs associating trauma scores with mortality, or predictive power using receiver operating characteristic analysis.

Capacity to predict mortality

Seventy articles reported on the relationship between ISS and mortality (online supplementary file 3). Of these, 43 articles (61%) reported ISS to be a good or strong predictor of death in their population samples.7 9 19 20 21 22–24 25 26 27 28 29 30–32 33 34 35 36 37 38 39 40 41 42 43 44 45–51 52 53 54 55 56 57 58 59 Fourteen articles reported ISS to be a weak predictor of mortality.10 12 14 60 61 62 63 64 65 66 67–69 Articles noting that ISS was not significantly associated with mortality had considerable missing data or very small sample sizes.71 72 Among the six articles evaluating multiple scoring systems, two articles reported that ISS was the weakest predictor of mortality compared with other measures evaluated.12 73 In two other articles, the relationship between mortality and ISS was implicitly accepted and reported that moderate or severe ISS was the reason for increased mortality.19 20 Of the seven articles that evaluated the use of AIS,13 four (57%) reported that the score was a significant or valid predictor of mortality.21 75–77 One study reported that AIS was strongly associated with hospital admission, and another confirmed a significant relationship between AIS score and injury severity.22 78

Among the 24 articles evaluating the predictive power of RTS (online supplementary file 3), 18 (75%) reported a significant inverse relationship with mortality.7 12 13 14 73 23–26 28 30 31 33 79 34 35 70 80 Another article reported that RTS was similar to AIS, KTS, and the South African Triage Scale when predicting in-hospital mortality.15 Of the 15 studies that evaluated the predictive performance of KTS 10 (68%) reported it to be among the main predictors of mortality.12 13–15 73 79 81 82 83 84 85 86 87 88 Two of these 10 studies (20%) noted that KTS equally predicted mortality compared with other scores, such as RTS, ISS, AIS, TRISS, and GCS.13 14 Other scores that were reported to be significant predictors of mortality in the context studied included SAPS II (Simplified Acute Physiology Score), PRISM (Pediatric Risk of Mortality), Physiologic Severity Score, and PTS (Pediatric Trauma Score).23 24 60 61

Of the 27 articles reporting TRISS assessments (online supplementary file 3), 9 studies indicated that the score provided a good estimation of survival probability,26–28 30 31 89 90 while 5 studies found it to be a moderate or weak predictor.14 73 74 75 91 At least 10 studies observed mortality that was greater than that predicted by TRISS.7 10 12 72 73 94 95 96 97 98 For example, a study in Iran predicted 3.1% mortality using the TRISS method and observed 8% mortality.99 Another study in India predicted 15.7% mortality, but observed 33.3% mortality among trauma patients.100 A prospective study in Uganda reported that 25 of 39 deaths were unexpected according to TRISS methodology, and that 24 of these deaths were deemed preventable by peer review.101 Only three articles (11%) reported that TRISS was a good predictor of mortality.14 26 27 Seven articles (25%) explicitly...
reported that the TRISS model did not accurately predict survival of trauma patients in their settings because the Major Trauma Outcome Study (MTOS) cohort was not the appropriate standard given the resources available.10,25,28,30,75,96

Feasibility and appropriateness in implementation

Five articles specifically stated that no trauma scoring was used, with one study noting that a lack of available information made any kind of physiological scoring impossible.97–101 These studies were conducted in Botswana, Burkina Faso, Ghana, Mexico, South Africa, and Vietnam. The authors pointed to specific challenges and constraints of trauma scoring that may contribute to inaccuracies in survival probability estimates and/or underestimation of injury severity. These barriers included a lack of equipment, human resources, and poor record keeping which in turn resulted in insufficient or poor-quality data (table 1).

Twelve articles commented on the limited feasibility of ISS and/or AIS due to limited technological capacity. That appropriate AIS coding requires diagnostic or radiographic procedures, such as accurate CT scans, was also a common constraint. One research group reported that patients referred to other procedures, such as accurate CT scans, was also a common constraint. That appropriate AIS coding requires diagnostic or radiographic procedures needed for accurate coding, resulting in potentially underestimated injury severity.102 Assessing ISS may be also impossible due to a patient’s inability to pay out-of-pocket costs associated with sophisticated investigations.33 A study in Nigeria compared RTS classification with panel consensus on preventable deaths, finding that the different processes generated different conclusions: RTS-based classification determined that 81% of deaths were preventable, whereas panel consensus identified 22.1% of deaths as preventable.104

Usage in QI efforts

Although few studies recounted how trauma scores had been used in QI processes several QI-related outcomes were reported to be positively correlated to key trauma scores (table 2).109–111 One study, focused on hospital-based trauma QI in Pakistan, compares ISS, GCS, and RTS scores before and after the implementation of QI initiatives as a means of demonstrating clinical efficacy.31 A study in Nigeria compared RTS classification of injury severity with panel consensus on preventable deaths.

Table 1  Reported feasibility of trauma scoring systems

<table>
<thead>
<tr>
<th>Trauma score</th>
<th>Opportunities</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>AIS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RTS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>KTS</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>TRISS</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

AIS, Abbreviated Injury Scale; eISS, estimated ISS; ISS, Injury Severity Score; KTS, Kampala Trauma Score; RTS, Revised Trauma Score; TRISS, Trauma and Injury Severity Score.

Table 2  Additional reported correlates of trauma scores

<table>
<thead>
<tr>
<th>Trauma score</th>
<th>Positive correlates</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>Incidence of organ dysfunction</td>
</tr>
<tr>
<td></td>
<td>Length of stay/prolonged or continued hospitalization</td>
</tr>
<tr>
<td></td>
<td>Morbidity</td>
</tr>
<tr>
<td></td>
<td>Postoperative complications</td>
</tr>
<tr>
<td></td>
<td>Severity/grades of anatomic-specific injuries</td>
</tr>
<tr>
<td>AIS</td>
<td>Hospital admission</td>
</tr>
<tr>
<td></td>
<td>Long-term functional outcomes</td>
</tr>
<tr>
<td></td>
<td>Severity/grades of anatomic-specific injuries</td>
</tr>
<tr>
<td>RTS</td>
<td>Hospital admission</td>
</tr>
<tr>
<td></td>
<td>Length of stay/prolonged or continued hospitalization</td>
</tr>
<tr>
<td></td>
<td>Morbidity</td>
</tr>
<tr>
<td></td>
<td>Nosocomial infections</td>
</tr>
<tr>
<td></td>
<td>Postoperative complications</td>
</tr>
<tr>
<td>KTS</td>
<td>Hospital admission</td>
</tr>
<tr>
<td></td>
<td>Length of stay/prolonged or continued hospitalization</td>
</tr>
<tr>
<td></td>
<td>Prehospital transportation</td>
</tr>
<tr>
<td>TRISS</td>
<td>Length of stay/prolonged or continued hospitalization</td>
</tr>
<tr>
<td></td>
<td>Nosocomial infections</td>
</tr>
</tbody>
</table>

*Three studies reported KTS was a significant predictor of need for admission; one study reported KTS was poor in predicting need for admission.96
†Three studies reported KTS predicted longer hospitalization; one study reported KTS was not a good indicator of prolonged hospitalization.75
AIS, Abbreviated Injury Scale; ISS, Injury Severity Score; KTS, Kampala Trauma Score; RTS, Revised Trauma Score; TRISS, Trauma and Injury Severity Score.
Using ISS to examine mortality patterns in the USA and Mexico, Arreola-Risa et al concluded that improvements in intensive care and other hospital-based technologies would be less efficacious in reducing overall trauma mortality than improvements in prevention, prehospital transport and care, and emergency room care. In Uganda, QI processes using TRISS methodology to identify unexpected deaths resulted in 25 cases subject to peer reviews. This process identified critical errors in care management, including missed diagnoses, inadequate intravenous therapy, delayed surgery, and poor airway control. Because more than half of reviewed deaths were due to missed injuries, peers identified clinical evaluation and investigation as well as interpretation of radiological signs as key areas for improvement. In their TRISS analysis, Zargar et al partially attributed unsatisfactory quality of care to insufficient usage of ambulances, delayed transportation to hospitals, frequent referrals between facilities, and inadequate resuscitation by emergency services.

**DISCUSSION**

This review reveals that GCS and ISS remain the prominent trauma scoring systems in LMICs. GCS may be often used because of its ease of implementation and because it remains a strong predictor of mortality. However, its focus on neurological state, though important, may not be sufficiently holistic to assess injury severity. It may assign low scores to injuries without immediate neurological effects that are, nonetheless, severe and cause mortality over time. This could pose problems for use in prehospital care or triage procedures.

ISS was generally a good to strong predictor of mortality. However, ISS was found to be the weakest relative to other trauma scoring systems, including RTS, KTS, NISS (New Injury Severity Score), and TRISS. Persistent challenges faced in implementation suggest that the computation of ISS in LMIC settings is not always accurate. AIS, and subsequently ISS, are limited in that they do not consider the effect of age and exclude all but the most serious injury in one body region. Studies noted that conservative estimates were made when appropriate resources were not available for complete assessment. A wide variety of injury combinations may correspond to different injury severity in reality, but generate similar ISS values. Moreover, the lack of imaging capacity may be a critical limitation to achieving more accurate ISS assessments in LMIC settings. Despite these issues, it remains the most commonly used score and is often implemented without validated adaptations for resource-constrained settings.

KTS has been highlighted in a number of articles as a viable, validated alternative that predicts mortality similarly to RTS and ISS. The authors cautioned its premature integration into clinical decision-making, but emphasized its capacity to discriminate injury severity using a minimal data set and relatively simple settings. Continued rigorous evaluation of KTS as a triage tool and its ability to discriminate mortality may bolster confidence in its validity. The failure of TRISS to accurately predict survival in retrospective studies suggests that facilities do not generally collect the data necessary to be effective. Furthermore, MTOS, the retrospective study of trauma epidemiology and outcomes conducted in the USA used by TRISS computations, may be an inappropriate standard to use in resource-constrained settings. Modified TRISS coefficients generated from MTOS-like evaluations in LMICs are needed to improve injury severity assessments using this scoring system in these settings.

In several cases, limited human resources led to missing or poor-quality recording of patient characteristics and conditions. Sufficient staff is required for appropriate recording of data, such as vital signs, and consistent administration of a trauma database is time consuming and can require dedicated staff. General recommendations included periodically reviewing collected data to improve the quality of trauma management and establish continuous surveillance at facilities. Plans to analyze injury patterns in LMICs should first examine the completeness of data sources and consider adjusting analyses for such resource constraints. Establishing sustainable trauma databases using local resources will also require finding key champions to promote culture change at an institutional level and impress the value of data collection among hospital personnel. The situation thus emphasizes the importance of health systems strengthening initiatives focused on capacity building to provide appropriate training. Such initiatives can have both short and long-term impacts in trauma care for appropriate triage, monitoring, and QI.

Trauma scoring methods are key to QI programs. Though several studies noted that QI measures should be taken, trauma scores were not readily used in formal QI processes. This may be a missed opportunity as findings from a few studies suggest trauma scores may be useful tools for identifying preventable deaths and evaluating care processes in LMICs—measures that have the potential to generate context-appropriate interventions for trauma care improvement. Yet, barriers to change, such as cultural norms, education and training, or inadequate resources, may exist and need to be further explored through implementation research. Capitalizing on existing trauma scoring systems may be an efficient use of resources in developing trauma care in these settings.

Interpretations of this review should take into account a number of limitations. Its broad scope highlighted which scoring systems have been applied in LMICs and generally reports on their appropriateness across a range of populations and types of injuries. Understanding trauma scoring for specific populations, such as children or specific injuries, will require further disaggregating findings. Publication bias may also play a significant role. To determine which scoring systems are relevant to LMICs, it will be important to understand the shortcomings of particular trauma scores and which can most negatively affect the effectiveness and efficiency of care processes.

Lesser used scores may prove feasible and appropriate for LMICs though they remain understudied. Among prospective studies, it was difficult to discern whether reported systems were used regularly or only for the purposes of research. Identifying which systems remain in regular use beyond organized studies would better represent prevalence and feasibility of existing systems. Specifying whether trauma scoring systems were used nationally or by particular areas or regions of a country could highlight features that affect the implementation of scoring systems or other interventions.

**CONCLUSIONS**

Trauma care in LMICs has largely been guided by advancements of HICs in recent decades. Trauma scoring is a key tool across stages of trauma care and beyond, used for quality assurance and improvement, clinical governance, resource allocation, and research. Implementing systems designed for HICs may not be immediately relevant to the burden and resources available in LMICs. In multiple instances, adaptations or alternative scoring systems have been more appropriate in assessing trauma in such...
settings. Yet, the use of these systems remains limited. Future research should focus on determining the optimal trauma scoring system under resource constraints. As a tool to accurately identify the most severe cases, system bottlenecks, and preventable deaths, trauma scoring remains an underused means to informing care for the distinct burden, resources, and infrastructure available in LMIC settings.

**Contributors** CJ and RD conceived the study objectives and oversaw overall study design and implementation. IF and GW designed the search strategy and data acquisition procedures. IF, GW, II, ZM, and MC implemented the screening of articles for eligibility for review as well as extracted data from relevant articles. IF with support from MC interpreted and synthesized the data as well as drafted the article. All authors reviewed and approved the final article.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is eligible for review as well as extracted data from relevant articles. IF with support from MC interpreted and synthesized the data as well as drafted the article. All authors reviewed and approved the final article.

**COMPETING INTERESTS**

None declared.

**OPEN ACCESS**

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.


