Quality Indicators for Evaluating Trauma Care

A Scoping Review

Henry Thomas Stelfox, MD, PhD; Barbara Bobranska-Artiuch, MD; Avery Nathens, MD, PhD; Sharon E. Straus, MD, MSc

Objectives: To systematically review the literature on quality indicators (QIs) for evaluating trauma care, identify QIs, map their definitions, and examine the evidence base in support of the QIs.

Data Sources: We searched MEDLINE, EMBASE, CINAHL, Cochrane Database of Systematic Reviews, Cochrane Database of Abstracts of Reviews of Effects, and Cochrane Central Register of Controlled Trials from the earliest available date through January 14, 2009. To increase the sensitivity of the search, we also searched the grey literature and select journals by hand, reviewed reference lists to identify additional studies, and contacted experts in the field.

Study Selection and Data Extraction: We selected all articles that identified or proposed 1 or more QIs to evaluate the quality of care delivered to patients with major traumatic injuries. Minimum inclusion criteria were a description of 1 or more QIs designed to evaluate patients with major traumatic injuries (defined as multi-system injuries resulting in hospitalization or death) and focused on prehospital care, hospital care, posthospital care, or secondary injury prevention.

Data Synthesis: The literature search identified 6869 citations. Review of abstracts led to the retrieval of 538 full-text articles for assessment, of which 192 articles were selected for review. Of these, 128 (66.7%) articles were original research, predominantly trauma database case series (57 [29.7%]) and cohort studies (55 [28.6%]), whereas 37 (19.3%) were narrative reviews and 8 (4.2%) were guidelines. A total of 1572 QIs in trauma care were identified and classified into 8 categories: non-American College of Surgeons Committee on Trauma (ACS-COT) audit filters (42.0%), ACS-COT audit filters (19.1%), patient safety indicators (13.2%), trauma center/system criteria (10.2%), indicators measuring or benchmarking outcomes of care (7.4%), peer review (5.5%), general audit measures (1.8%), and guideline availability or adherence (0.8%). Measures of prehospital and hospital processes (60.4%) and outcomes (22.8%) were the most common QIs identified. Posthospital and secondary injury prevention QIs accounted for less than 5% of QIs.

Conclusions: Many QIs for evaluating the quality of trauma care have been proposed, but the evidence to support these indicators is not strong. Practical recommendations to select QIs to measure the quality of trauma care will require systematic reviews of identified candidate indicators and empirical studies to fill the knowledge gaps for postacute QIs.


I

JURY IS ONE OF THE LEADING causes of death in almost every country in the world. Each year, injuries affect 700 million people around the world and result in more than 5 million deaths.1 However, deaths are only the tip of the injury iceberg. In the United States alone, injured Americans annually make 30 million emergency department visits resulting in 1.9 million hospital admissions at direct medical costs of approximately $80 billion.2 The human and societal burden is even greater, with many survivors never returning to school, work, or their “regular” lives.3,4

Surgical services provide injured patients with pivotal treatment for what is a major cause of morbidity and mortality. However, there is growing evidence that the best treatments and strategies for these patients are not always implemented.5,6 Several studies have documented that as many as half of all patients with major traumatic injuries do not receive recommended care.7-11 In addition, medical errors are common among critically ill trauma patients.12,13 Review of trauma deaths in hospitals has suggested that between 2.5% and 14% of medical errors are preventable.14-19 As a result, the outcomes of patients with major traumatic injuries are not as good as they could be with better translation of the best research evidence at the bedside.

To improve trauma care, the quality of care first needs to be measured using evidence-based tools. However, it is unclear
if the proper tools have yet been developed. The purpose of this scoping review was to systematically review the literature on quality indicators (QIs) in evaluating trauma care, identify QIs, map their definitions, and examine the evidence base in support of QIs.

### METHODS

#### SEARCH STRATEGY

Relevant articles were identified by searching the following databases on January 14, 2009, from the first date available: Ovid MEDLINE; Ovid EMBASE; Ovid CINAHL; and Cochrane Database of Systematic Reviews, Cochrane Database of Abstracts of Reviews of Effects, and Cochrane Central Register of Controlled Trials. Searches were performed with no language of publication restrictions. Combinations of the following search terms were used: trauma, injury, quality indicator, quality assurance, quality control, performance improvement, quality measure, best practice, and audit. The Cochrane Library was searched separately using the search term trauma. To increase the sensitivity of the search strategy, we also searched the grey literature. This search included identifying and searching Web sites of relevant trauma organizations (American College of Surgeons, American Association for the Surgery of Trauma, Eastern Association for the Surgery of Trauma, American Trauma Society, British Trauma Society, Trauma Association of Canada, Australasian Trauma Society, Orthopedic Trauma Association, Western Trauma Association, Trauma.org, The Society of Trauma Nurses, International Trauma Anaesthesia, and Critical Care Society) and Google using combinations of the following search terms: trauma, injury, quality indicator, quality of care, quality assurance, and quality control. In addition, a query was sent to 2 medical librarian listservs (CANMED-LIB and MEDLIB) asking about relevant materials, and responses were collected for 1 week. Appropriate wildcards were used in all searches to account for plural words and variations in spelling. Additional articles were identified by searching the bibliographies of those articles identified in the searches and contacting experts in the field. Selected journals (Journal of Trauma: Injury, Infection and Critical Care, Injury: International Journal of the Care of the Injured, Annals of Surgery, Surgery, and Journal of the American College of Surgeons) were manually searched from August 1, 2006, through July 31, 2007, to ensure that important articles were not missed.

#### ARTICLE SELECTION

We selected all articles that identified or proposed 1 or more QIs for evaluating the quality of trauma care. We included all articles, both research and nonresearch. For this study, a QI was defined as a performance measure that compared actual care against ideal criteria, a tool to help assess quality of care. 20 Minimum inclusion criteria were (1) the article included at least 1 QI; (2) the QI was designed to evaluate the care of patients with major traumatic injuries; and (3) the QI focused on trauma care specific to at least 1 of the following phases of patient care: prehospital, hospital, posthospital, or secondary injury prevention. Injury was defined as the “transfer of energy applied clinically.” 21 A major injury was defined as an injury resulting in at least 1 of the following: a multisystem injury (involving ≥2 regions of the body), an Injury Severity Score higher than 9, patient hospitalization, or patient death. 22,23 We did not include studies of patients with injuries secondary to poisonings, adverse effects of drugs or biological substances, and isolated hip fractures. 22,24 These conservative criteria were chosen because we wanted to optimize our chances for identifying relevant QIs.

#### RESULTS

The literature search identified 6869 articles. Review of abstracts led to retrieval of 402 full-text articles for assessment plus 72 articles from review of references and 64 articles from the hand search. We identified 192 ar-

### Table 1. Quality Indicator Classification Scheme

<table>
<thead>
<tr>
<th>Phase of Care</th>
<th>Structure</th>
<th>Process</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prehospital</td>
<td>Field triage</td>
<td>EMS response time</td>
<td>Prehospital deaths</td>
</tr>
<tr>
<td>Hospital</td>
<td>24-h CT scanner</td>
<td>ED time</td>
<td>Hospital deaths</td>
</tr>
<tr>
<td>Posthospital</td>
<td>Protocol for rehabilitation referral</td>
<td>Rehabilitation facility wait time</td>
<td>Admission to long-term care</td>
</tr>
<tr>
<td>Secondary prevention</td>
<td>Alcohol screening and brief intervention program</td>
<td>Alcohol readmission</td>
<td>Recurrent injuries</td>
</tr>
</tbody>
</table>

Abbreviations: CT, computed tomography; ED, emergency department; EMS, emergency medical services.

*Table data are sample quality indicators identified in review.*
articles written in 3 languages (English, German, and Spanish) for inclusion in the study (Figure 1).* The most common reason for excluding articles after full-text review was the absence of a QI.

DESCRIPTION OF THE ARTICLES

Table 2 summarizes the characteristics of the articles (see eTable for descriptions of individual articles; http://www.archsurg.com). Most articles were research studies (66.7%) primarily consisting of case series (29.7%), cohort studies (28.6%), cross-sectional studies (4.2%), and nonrandomized controlled studies (2.6%) using established trauma databases (Table 2). The non–original research articles (33.3%) consisted primarily of narrative review articles (19.3%). Most articles were published in the United States (57.8%) in the last 2 decades. For most articles, the academic status of the corresponding institution was unclear. Few were supported by grants, primarily from the government (14.6%) or private foundations (7.3%).

Most articles focused on QIs to evaluate the quality of trauma care for adult patients, whereas few focused exclusively on pediatric patients (7.8%). The mecha-
DESCRIPTION OF THE QIs

Table 3 summarizes the characteristics of QIs in trauma care. We identified 1572 QIs in the 192 articles, a median of 3 QIs per article (interquartile range, 1-11) with a range of 1 to 60 QIs per article. The QIs were described in the articles using 103 different terms (eg, QI, audit filter, standards for trauma care, safety standard, annual feedback, avoidable factors, criteria deficiencies, etc). The most common source of QIs were authors of articles or local health care provider groups (56.9%), followed by the American College of Surgeons Committee on Trauma (ACSCOT) (26.5%) (Table 3). We reviewed 8 components of QI definitions. Virtually all QIs had a descriptive statement (1568 of 1572 [99.7%]), but only a few provided a list of data elements, specifications for data element collection, a description of the population and/or analytic models used to construct the measure, timing of data collection, the format for presentation of the results, and the timing of reporting. The data sources for measuring QIs were specified for 942 indicators (59.9%), with trauma registries (27.2%) and medical record review (24.6%) being the most common sources. After reviewing the definitions and data sources provided for the QIs, the 2 reviewers determined that 959 (61.0%) of the QIs were potentially operational.

Table 4 categorizes the QIs based on content matter. The 2 most frequent categories were non-ACSCOT (42.0%) and ACSCOT (19.1%) audit filters. Some QIs related to guideline availability or adherence were identified. Figure 2 summarizes the distribution of the QIs according to our prespecified conceptual model. Measures of prehospital and hospital processes of care were the most common QIs identified. Posthospital and secondary injury prevention QIs accounted for less than 5% of QIs.

IMPLEMENTATION OF QUALITY INDICATORS

Table 5 summarizes author recommendations regarding use of QIs in trauma care. Authors of the articles recommended that 616 QIs (39.2%) be imple-
mented and used routinely, whereas 78 QIs (5.0%) should not be used. The authors recommended that the QIs be used for quality evaluation in 377 cases (24.0%) and as a quality improvement tool in 308 cases (19.6%). Additional research was recommended for 94 QIs (6.0%) before their implementation in clinical practice. Table 6 lists the 10 QIs identified most frequently in original research studies, which are potential candidates for systematic review and further evaluation.

### Table 6. Candidate Quality Indicators (QIs) for Systematic Review

<table>
<thead>
<tr>
<th>QI</th>
<th>Types of Original Research Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peer review of trauma deaths to evaluate quality of care and determine whether the death was potentially preventable</td>
<td>19 Case series24,40,41,50,52,58,60,61,75,77,95,97,119,127,128,148,161,169,186,190,200; 18 cohort studies24,40,42,43,50,52,65,77,80,91,119,127,128,148,161,169,186,190,200; 2 nonrandomized controlled trials16,205; and 1 cross-sectional study17</td>
</tr>
<tr>
<td>Hospital mortalitya</td>
<td>17 Cohort studies24,38,40,41,50,52,65,77,80,91,119,127,128,148,161,169,186,190,200,206; 9 case series42,55,57,77,80,91,128,139,140,146,161,186,190,200; 4 before-and-after case series102,130,157,166; 3 nonrandomized controlled studies103,203,206; 2 cross-sectional studies22,104; and 1 case-control studya</td>
</tr>
<tr>
<td>Complications during hospital stayb</td>
<td>9 Cohort studies23,48,50,52,65,77,128,139,140,146,186,190; 5 case series12,24,30,158,159; 4 cross-sectional studies12,24,30,158,159; 2 before-and-after case series95,53; 1 case-control study4; and 1 consensus method17</td>
</tr>
<tr>
<td>Patient treated at the scene longer than X min (range, 10-30 min)</td>
<td>10 Cohort studies23,48,50,52,65,77,80,91,128,139,140,146,186,190; 2 before-and-after case series74,171; 2 case series52,123; and 1 consensus method154</td>
</tr>
<tr>
<td>Glasgow Coma Scale score &lt;X (range, 9-14) and no CT scan of the head within X h (range, 1-4 h) of arrival</td>
<td>5 Cohort studies17,60,128,189; 3 before-and-after case series57,140,158; 3 case series42,50,208; 1 case-control study4; 1 cross-sectional study14; and 1 consensus method127</td>
</tr>
<tr>
<td>Time from patient hospital arrival to emergency surgical treatment (range, 24-48 h) of initial procedure</td>
<td>5 Cohort studies18,50,128,190; 3 before-and-after case series119,128,190; 1 nonrandomized controlled trial19; 1 case-control study1; 1 cross-sectional survey15; and 1 consensus method17</td>
</tr>
<tr>
<td>Unscheduled surgical treatment within X h (range, 24-48 h) of initial procedure</td>
<td>3 Cohort studies17,128,3, before-and-after case series34,157,159; 3 case series42,50,129; 1 case-control study4; and 1 consensus method17</td>
</tr>
<tr>
<td>Missed injuries, ie, injuries diagnosed/documented X h (range, 24 h to discharge) after admission</td>
<td>5 Cohort studies17,128,161; 3 case series24,55,152; 1 before-and-after case series158; and 1 consensus method17</td>
</tr>
<tr>
<td>Glasgow Coma Scale score &lt;X (range, 8-10) and airway not secured within X h (range, &lt;5 min) to before patient leaves the ED</td>
<td>3 Cohort studies3,128,190; 3 before-and-after case series57,140,158; and 2 case series92,208</td>
</tr>
</tbody>
</table>

a Percentages may not total 100 because of rounding.

In our review of the current literature on QIs to evaluate the quality of trauma care, we identified 192 articles and 1572 QIs. A large body of literature exists for evaluating trauma care and may be helpful for guiding development of evidence-based QIs, although few indicators appear to be explicitly defined and supported by valid criteria for collection. The articles available are predominantly focused on prehospital and hospital processes of care for adult populations and provide opportunities for systematic reviews. Very few articles were identified that focused on QIs for evaluating the quality of trauma care for children or postacute care, suggesting a need for empirical research in these domains.

The single most important result from our review is that we did not find a common set of clearly defined, evidence-based, broadly accepted QIs for evaluating the quality of trauma care. Rather, our review identified a large group of heterogeneous indicators that could be broadly categorized into 8 groups. It is not possible to have a transparent, explicit, systematic, data-driven performance measurement feedback system if there are no generally agreed upon measures. How can we explain the contrast between our results and the interest and resources that are being directed toward measuring quality around the globe? Despite its theoretical appeal, developing quality measures is challenging. One, physicians, researchers, and administrators in a field have to communicate using the same language. We identified 103 terms used to identify or describe QIs in our review. Although the field of quality improvement is fundamentally a process of change...
in human behavior and is driven largely by experiential learning, it would be nearly impossible to effectively communicate with such a large variation in common terms. An agreement on a basic taxonomy is critical similar to the recently published SQUIRE (Standards for QUality Improvement Reporting Excellence) guidelines. Two, trauma care has until recently been intensively local, with quality improvement based on local surgical audits. Quality measures should be developed to serve national goals, but should be implemented locally. Our review highlights the frequency with which local QIs have been developed. Three, performance measurement is facilitated with a large, high-quality evidence base. For example, the American College of Cardiology and the American Heart Association have jointly published 11 evidence-based QIs for adults with myocardial infarction. These indicators are widely accepted because the underlying data arise from large, multicenter, randomized, controlled trials. Conversely, physicians caring for patients with traumatic injuries do not have the same evidence base to draw upon in developing QIs. Four, developing QIs for diseases or disorders that are homogeneous is easier than for those that are heterogeneous. Trauma is not a disease but a collection of injuries, and QIs may therefore need to focus on common pathways and outcomes.

Just because a common set of evidence-based QIs of trauma care have not been developed does not mean that they cannot be developed. Recognizing this challenge, the American College of Surgeons has proposed the Trauma Quality Improvement Project, modeled after the successful National Surgical Quality Improvement Program. A validated, risk-adjusted, outcomes-based program to measure and improve the quality of surgical care. The Trauma Quality Improvement Project is a proposal to measure risk-adjusted performance of trauma centers, identify institutional characteristics associated with good patient outcomes, and promote those institutional characteristics among trauma organizations. A feedback mechanism involving confidential report cards would be used to inform each trauma center of its performance. This proposal is potentially an important step in improving the quality of trauma care because it may provide for the first systematic measuring and reporting of trauma care quality. However, more is needed. Valid measures of structure and process of care are also important because they provide a direct means for programs to evaluate themselves. Finally, outcome measures other than hospital mortality will be needed to capture additional outcome dimensions such as health-related quality-of-life and to provide greater specificity for different components of patient care. Our review begins to outline where sufficient literature exists to evaluate candidate indicators (Table 6) and where gaps exist (Figure 2) and empirical research is needed.

There are limitations to this review. First, our search may not have been exhaustive, despite the search of multiple databases using comprehensive search strategies with the assistance of an information specialist and imposing no language restrictions. Nevertheless, it is unlikely that our search missed broad categories of important QIs. Second, it is difficult to extract accurate data from all publications. Some articles are difficult to obtain (3 in our study), some do not disclose all materials or methods used, and results are often unclear and difficult to interpret. Third, categorizing articles by type, even with the assistance of predefined data abstraction tools and classification schemes, is partly subjective.

CONCLUSIONS

We have shown that many QIs for evaluating the quality of trauma care have been proposed but that the distribution of these indicators is skewed across different components of health care quality and trauma systems. Our results should promote future research in 3 areas. First, systematic reviews are warranted to explore whether evidence-based QIs can be developed from the existing literature for acute-care measures. Second, empirical research is needed to fill the knowledge gaps for postacute QIs. Third, research is needed on how QIs are currently used by trauma systems and centers to measure and improve trauma care as a means to better understanding how to effectively incorporate newly developed indicators into performance-improvement initiatives. Identifying, developing, evaluating, and implementing evidence-based QIs is essential to measuring and improving trauma care.

Accepted for Publication: July 23, 2009.

Correspondence: Henry Thomas Stelfox, MD, PhD, Department of Critical Care Medicine, University of Calgary, South Tower, Room 1105, 1403-29 St NW, Calgary, Alberta T2N 2T9, Canada (Tom.Stelfox@albertahealthservices.ca).

Author Contributions: Drs Stelfox and Straus had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Stelfox, Nathens, and Straus. Acquisition of data: Stelfox, Bobranska-Artiuch, and Straus. Analysis and interpretation of data: Stelfox, Bobranska-Artiuch, Nathens, and Straus. Draft of the manuscript: Stelfox. Critical revision of the manuscript: Stelfox, Bobranska-Artiuch, Nathens, and Straus. Statistical analysis: Stelfox. Administrative, technical, and material support: Stelfox and Straus. Study supervision: Stelfox.

Financial Disclosure: None reported.

Funding/Support: The project was supported by Synthesis Grant KRS-91770 from the Canadian Institutes of Health Research. Dr Stelfox is supported by a New Investigator award from the Canadian Institutes of Health Research.

Role of the Sponsors: The funding sources had no role in the design, conduct, or reporting of this study.

Online-Only Material: The eTable is available at http://www.archsurg.com.

Additional Contributions: Laure Perrier, MS, assisted with the literature search strategy, and Kelly Mrklias, MSc, assisted with translation of Spanish and German language articles.

REFERENCES


INVITED CRITIQUE

What Makes a “Good” Quality Indicator?

In their study, Stelfox and colleagues conduct a literature review to identify QIs for trauma. The authors should be commended for their exhaustive literature review and valuable contribution to the literature. However, the present article lacks a definition of what makes a “good” QI. This lack of a standard for evaluating measures limits the conclusions that can draw from this study. Although the authors close with a set of candidate measures, they are unable to describe why these measures are selected, other than a gestalt that they were mentioned most frequently.

What makes a good QI? The National Quality Forum (http://www.qualityforum.org/), a leading organization for the endorsement of QIs, considers several criteria when evaluating a new measure:

**Importance:** The QI must be relevant to a large number of patients, involve a high-risk condition, or in some other way represent a large opportunity for improvement.

**Scientific acceptability:** The measure must be reliable and valid. Reliability means the indicator gives the same result on repeated measures, and validity implies that the indicator measures what it is intended to measure.

Feasibility: Data for reporting QIs must be feasible to obtain. Some QIs are readily available; others require extensive data collection from medical records. Measures based on data that are difficult to obtain must be extremely valuable or they will result in misspent resources.

Usability: The results of any measure must be understood by the intended audiences (eg, consumers, health care providers, and payers). Measures that are difficult to understand will not be translated into meaningful improvement.

As Stelfox and colleagues move forward, perhaps their next step could be determining which of the many QIs identified in their literature review are the most important, valid, feasible, and usable.

Justin B. Dimick, MD, MPH

**Author Affiliation:** Department of Surgery, University of Michigan, Ann Arbor.

**Correspondence:** Dr Dimick, Department of Surgery, University of Michigan, 211 N Fourth Ave, Ste 301, Ann Arbor, MI 48104 (jdimick@med.umich.edu).

**Financial Disclosure:** None reported.